Chapter 3
The Role of Social Media Geographic Information (SMGI) in Spatial Planning

Michele Campagna, Roberta Floris, Pierangelo Massa, Anastacia Girsheva and Konstantin Ivanov

Abstract This contribution reports on ongoing research carried on by the authors on the role of Social Media Geographic Information in spatial planning, design, and decision-making. Explicit and Implicit Volunteered Geographic Information (VGI) from social media platforms, namely Social Media Geographic Information (SMGI) resources, were used to explore novel methods and tools for analysis and knowledge construction. The results concern three main research streams carried on with the common feature of integrating social media and other volunteered and authoritative sources of information from Spatial Data Infrastructures (SDI). These findings demonstrated that the integration of SMGI with more traditional Authoritative Geographic Information (A-GI) may offer a high potential for eliciting pluralist knowledge for spatial planning.

M. Campagna (✉) · R. Floris · P. Massa
DICAAR, Università Di Cagliari, Via Marengo 2, 09123 Cagliari, Italy
e-mail: campagna@unica.it

R. Floris
e-mail: roberta.floris@unica.it

P. Massa
e-mail: pmassa@unica.it

A. Girsheva · K. Ivanov
Università Di Cagliari, Via Marengo 2, 09123 Cagliari, Italy
e-mail: grisheva@unica.it

K. Ivanov
e-mail: konstantin.ivn@gmail.com

K. Ivanov
Tomsk Polytechnic University, 30 Lenin Avenue, Tomsk, Russia 634050

© Springer International Publishing Switzerland 2015

S. Geertman et al. (eds.), *Planning Support Systems and Smart Cities*,
Lecture Notes in Geoinformation and Cartography,
DOI 10.1007/978-3-319-18368-8_3

1 Introduction

Since last decade, advances in the Information and Communication Technologies (ICT), the Internet, and more recently, in the Web 2.0 technologies are increasingly channeling digital Geographic Information (GI) into daily life of a growing number of users. This wealth of GI may foster notable innovations in spatial planning methodologies and practices, for the majority of information required to support analysis, design, and decision is inherently spatial in nature. However, this hypothesis should be carefully tested and much work is still needed to develop methods and tools capable to offer planners reliable and user-friendly methods and tools.

Since the late 1990s, developments in Spatial Data Infrastructures (SDI) granted the access to digital data, produced and maintained by public or private organizations for institutional or business purposes. In Europe, the Directive 2007/02/CE establishing a shared INfrastructure for SPatial InfoRmation (INSPIRE), is leading to the development of National and Regional SDIs in the Member States according to common data, technology, and shared standards, so enabling public access and reuse of available official spatial data, or Authoritative Geographic Information (A-GI). The term 'authoritative' refers to spatial data produced by experts, professionals, organizations, and mapping agencies for a mission under institutional or legal frameworks (Ball 2010; Goodchild and Glennon 2010). The production of A-GI by highly trained experts complies with specific requirements and quality assurance procedures, in order to guarantee accuracy and quality standards (Goodchild and Glennon, ibidem; Elwood et al. 2012). Furthermore, the authoritativeness of A-GI is assured by metadata, which describe content, quality, accuracy, authorship, conditions of use and other characteristics of this information (Nogueras-Iso et al. 2004).

At the same time, many platforms continue to flourish online thanks to continuous advances in Web 2.0 technologies, which enabled the production, collection, and diffusion of user-generated contents (Krumm et al. 2008), wherein the community plays a more fundamental role in data production (Bruns 2006). Hence the popularity of the term Volunteered Geographic Information (VGI) (Goodchild 2007), which refers to the user-generated contents with a geospatial component created by citizens acting as volunteer sensors. The concept of VGI encompasses a wide range of activities and practices, which may provide pluralist sources of both experiential knowledge from local communities and expert knowledge from professionals, generating unprecedented opportunities for enhancing democratic decision-making in spatial planning processes. In several countries worldwide, the use of VGI has been proven useful in many application domains such as emergency management (Zook et al. 2010), environmental monitoring, spatial planning (Poser and Dransch 2010), crisis management (Roche et al. 2013), as well as participatory processes within Citizen Science initiatives (Haklay 2013).

In addition, widespread popularity of social media platforms is fostering the diffusion of geo-referenced multimedia (Sui et al. 2013), or Social Media

Geographic Information (SMGI). The latter information sources may be easily accessed and shared by users, which seamlessly become producers and consumers of personal geo-referenced contents on location-aware social networks. SMGI represents a deviation from an early vision of VGI, inasmuch dissemination of geographic information is not the final purpose of production (Stefanidis et al. 2013). As a matter of fact, SMGI for its nature may be classified as implicit VGI, in contrast with explicit VGI, whose main purpose is the diffusion of geographical contents (Craglia et al. 2012).

Despite this distinction, SMGI could lead toward innovative scenarios for gathering and disseminating geographic information among million of users worldwide, eventually providing valuable insights about user perceptions or needs, opinions on places, daily-routine events, so helping to get better insights on local identities (Campagna 2014), flows of information, and social networking within societies (Stefanidis et al., ibidem).

However SMGI, unlike A-GI, due to its peculiar user-generated mode of production features Big Data characteristics (Caverlee 2010) and, consequently, traditional spatial analysis methodologies and techniques may be not fully adequate to take advantage of the enclosed knowledge potential. A possible way to address this challenge is given by computational social science, a new emerging discipline aiming at finding new methods and tools to tackle the complexity of Big Data management issues (Lazer et al. 2009). The management of geographic Big Data, their integration with A-GI, and the use of advanced analytics may enable the extraction of relevant knowledge to support decision-making in diverse fields including spatial planning and design. This approach might also inform smart city initiatives by supplying real-time dynamic pluralist knowledge on people's perception of places.

In the light of these premises, the authors present a critical review of their research findings on the integrated use of A-GI, VGI, and SMGI in the domain of spatial planning. The remainder of the chapter is organized as follows. In the next section, a brief discussion about the main components of smart city initiatives and the way such strategies could be affected by SMGI is given. In Sect. 3 the authors introduce a novel approach to SMGI analytics, proposing its application in three different case studies. Finally, Sect. 4 draws conclusions discussing the results and the relevance of SMGI for spatial planning.

2 Digital Geographic Information for Smart City Strategies

The wealth of digital geographic information about facts, opinions and concerns of users, made available by the Internet and Web 2.0 technologies, could affect current practices in spatial planning and smart city strategies offering the possibility for real-time monitoring of the needs and aspirations of local communities. Nowadays, the label 'smart city' identifies several strategies for dealing with problems generated by rapid urbanization and population growth in cities. In literature, several definitions of smart city can be found, providing a wide set of components that should be considered for the success of such strategies.

The Internet and Web 2.0 technologies play a central role to deal with several societal challenges, such as urban welfare, societal participation, environmental sustainability, and quality of life (Schaffers et al. 2010). Likewise, Information and Communication Technology (ICT) should be considered fundamental to integrate, connect, and make efficient the global system of infrastructures and services (Washburn and Sindhu 2010), or to improve livability and sustainability in the urban systems (Toppeta 2010). Technology is also fundamental to make 'smart cities' a source of spatial enablement for citizens, in order to improve access, sharing, and integration of spatial data with services (Roche et al. 2012). At the same time, technologies should allow innovative forms of communication, governance, and organization for the community engagement in evaluating and solving urban key problems (Batty et al. 2012). Therefore, several factors, such as governance, policies, and the community, enclosed in the political dimension, may substantially contribute to 'smart' development.

More specifically, local communities may play a critical role in the development of smart cities, due to the fact that these strategies directly affect the quality of citizens' life; hence, their needs and opinions play a major role for the transparency of such strategies. Therefore, SMGI may represent a valuable source of information regarding opinions, needs and perceptions of local communities, which could be used to inform 'smart city' strategies. However, the current lack of shared and reliable user-friendly methods for analysis and knowledge mining from SMGI could prevent to exploit the full potential from these sources. In order to address this issue, in the next section a novel approach for SMGI analytics is introduced and then discussed in the light of the results of three different case studies.

3 A Novel Approach to SMGI Analytics

The increasing SMGI production and availability over the web is paving the way to innovative analysis scenarios in spatial planning and geodesign, which in turn could be used to increase smart city performances. As introduced earlier, the integration of SMGI with A-GI may offer potentially boundless and affordable sources of information regarding not only geographic facts, but also perceptions, opinions, and feelings of local communities in space and time, or in other words it may help to depict the local identity or the genius-loci.

However, there is a lack of a shared analytical framework to collect, manage, and process this information for different purposes. In the next paragraphs, the authors present the findings of three projects which explore and formalize a novel approach to SMGI analytics, based upon the spatial, temporal, and textual analysis of SMGI, or STTx analytics. The first case study deals with the use of the geographic social media platform Place, I care! which enables the users to publish and interact with SMGI in a geo-browser, wherein the working space is an interactive map. Secondly, an original user-friendly tool for the collection and analysis of SMGI from social media is presented. Lastly, a wider case study demonstrates the role of SMGI in tourism planning.

3.1 Place, I Care! Crowdsourcing the Sense of Place and Supporting Community Dialogue

The first experience of the authors concerning VGI was developed back in 2011 in order to investigate how the opportunities generated by the innovation in social processes granted by new technologies could bring innovation in spatial planning and decision-making. After few attempts, trial and errors pilots using existing free tools, a brand new web application, called "Place, I care!" (PIC!), was built by the authors in order to fulfill the emerging requirements in a novel and integrated way. The results of the first pilots using PIC! were successful for framing a new SMGI analytics and developing the methods and tools presented in the following sections.

PIC! was originally designed as a geographic social networking platform to be used in urban and regional planning processes. The idea was to create a VGI planning support tool for collecting information from concerned citizens about the physical, environmental, and socio-cultural space, and for supporting the dialogues about urban and environmental issues, in a collaborative and participatory manner.

PIC! 1.0 is a web application which enables each user to easily create its own VGI projects in few steps, offering a flexibility that allows implementing a number of use-cases to fit a variety of working settings. While other similar platforms and applications were available on the web at the time of PIC design, no other was found which would allow creating as easily and flexibly many different use-cases not only to express and describe individual issues of concern or appreciation, but also to improve the possibility for collaborative discussion. This result was achieved integrating geo-web tools within a social network paradigm. Unlike in other geobrowsers and web mapping tools, in PIC! each post can be "Liked/Disliked" and commented by other users supporting the discussion.

Data collected with PIC! in several pilots enabled to understand "what and where" the main interests and concerns of the participants were. The more a post got comments or Like/Dislike, the more the issue was considered hot by the interacting community. If this is a common feature for a social network, the novelty here is the discussion is georeferenced and visualized in the map. As the number of post grows to tenths or hundreds it would be always possible to visualize what issues were important to the community, when, and where they are located, and to investigate patterns in the developing discourse through spatial analysis and statistics functions.

In summary, PIC! 1.0 was designed aiming at combining two main requirements: ease of use and availability to all, and robustness, which other geobrowsers could not guarantee to the final user. Hence, the main features of PIC! are the followings:

- project creation in few clicks
- flexible profiles and permissions management
- user-friendly multimedia posting
- like/dislike-ing and commenting
- advanced post query
- customizable layer management
- data export for the final user (to be implemented).

The implementation of the first PIC! pilots involving small groups between 60 and 100 participants produced only a limited amount of data (i.e. a few thousand georeferenced multimedia posts including point, line and polygon placemarks), but it immediately showed a great potential for discovering a new analytics which may be applied to much bigger data volumes. In Fig. 1, an example of a multimedia post with comments is shown in the PIC! interface.

From an analytical perspective, the integration of GIS and Social network data models allows one to discover more useful hints than would either A-GI or SMGI alone because (1) it integrates opinions and perceptions about facts and (2) it integrates the spatial and the thematic dimensions with time, multimedia (i.e. text, images, audio, and video), and the user, whose behavior eventually becomes a new dimension of analysis.

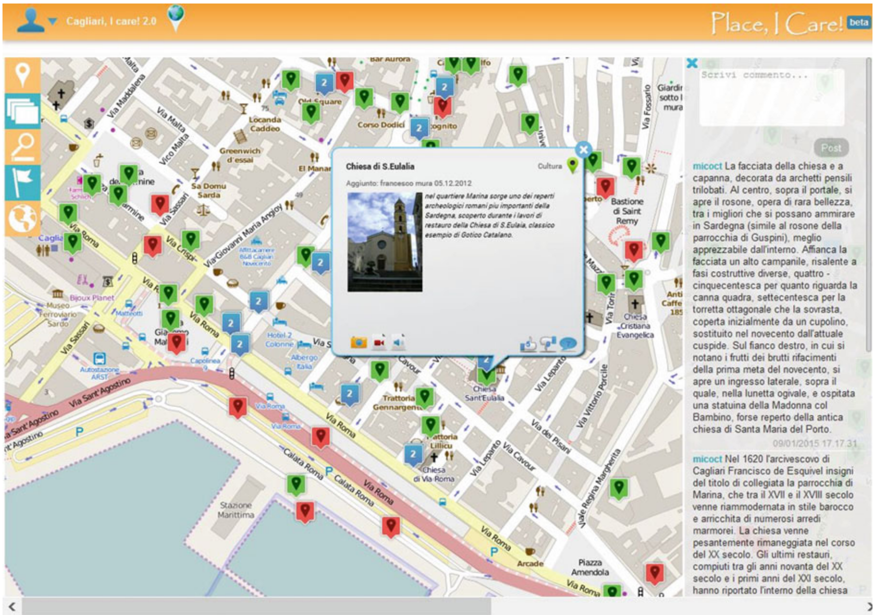

Fig. 1 Example of multimedia post in PIC!

The first exploratory analyses soon led to framing integrated analytics which, thanks to the specificity of the SMGI data model, includes:

- spatial analysis of user interests
- temporal analysis of users interests
- spatial statistics on user preferences
- multimedia content analysis on texts, images, audios, or videos
- user behavioral analysis
- combination of two or more of the previous such as the Spatio-Temporal-Textual Analysis (STTx) which enables to elicit what people discuss in space and time.

The analytic framework above shaped the design of the tools presented in Sect. 3.2, as well as, the study presented in Sect. 3.3.

3.2 SMGI: Spatext Tool

The early experiences with SMGI analytics informed the development of an original user-friendly tool, called Spatext, which enables one to extract information from multiple social media platforms and to seamlessly integrate it in a GIS environment for analysis. The tool provides several SMGI analysis methods including Spatial, Temporal and Textual (STTx) analysis. This SMGI Analytics suite is implemented as Python 2.7 add-in for ESRI ArcGIS©, offering a number of tools, which can be used mainly to (1) retrieve social media data from social media (including so far Twitter, YouTube, Wikimapia, and Instagram); (2) to geocode or georeference data; to carry on integrated (3) spatial, (4) temporal and (5) textual analyses. The number of analytical methods available in the tool is steadily increasing to include spatial-temporal clustering, in order to achieve user profiling, user movement analysis, and land use detection. Beside the desktop tool, Spatext analytical methods can also be used as Web Processing Services (WPS), in order to enable SMGI analytics via web interfaces. In Table 1 the analytical tools supplied by Spatext are classified by analytical functions.

The SMGI analytics application by Spatext is discussed through a number of examples conducted at the regional and local scale, which investigate both local communities' perceptions on relevant topics for spatial planning and the geography of places.

The first example concerns the analysis of georeferenced YouTube video metadata related to the term 'landscape' in Sardinia (Italy). The analysis of metadata spatial patterns shows that the term landscape in some Provinces (i.e. Cagliari and Olbia Tempio) is related to the coastal areas while in some other (i.e. Nuoro) concerns inner mountain areas, so expressing different local vocations. Similarly, in a second example, the spatial and textual analysis of video metadata contents in Cagliari (Italy) enabled the investigation of differences in perception regarding a number of neighborhoods by YouTube users. The results of the STTx define a clear

Table 1 Set of available tools in the Spatext suite

Spatext suite tools			
Harvesting SMGI from social media			
Tool name	Function	Textual query	Spatial query
Twitter extractor	Twitter SMGI extraction to table	Yes	No
YouTube extractor	YouTube SMGI extraction to feature class	Yes	Yes
Instagram extractor	Instagram SMGI extraction to feature class	No	Yes
Wikimapia extractor	Wikimapia SMGI extraction to feature class	Yes	Yes
Geocoding SMGI			
Tool name	Function	Manual	Automatic
Geocode address	Geocoding place/address from string	Yes	No
Geocode table	Batch-geocoding place/address from table	No	Yes
Georeferencing	Georeferencing SMGI coordinates by extractor (YouTube, Instagram, WikiMapia)	No	Yes
Textual analysis of SMGI			
Tool name	Function		
Attribute to string	Creation of a text file from an attribute field in a SMGI feature class for textual analysis		
Attribute to table	Creation of a table from an attribute field in a SMGI feature class for textual analysis		
Attribute to tag-cloud	Tag-clouding analysis from an attribute field in a SMGI feature class		
Selection to tag-cloud	Tag-clouding analysis from a spatial/attribute selection in a SMGI feature class		
Text to tag-cloud	Tag-clouding analysis from a text file		
Temporal analysis of SMGI			
Tool name	Function		
Identify month/weekday/day/hour	Add the month/weekday/day/hour of creation in a new field of SMGI feature class		
Trend day/hour	Creation of a 24 h/60 min time graph and statistic report from SMGI feature class		
Spatial analysis of SMGI			
Tool name	Function		
Decompose feature	Creation of a new feature class for each group in SMGI feature class (e.g. User)		
DBScan (density-based scan)	Run DBScan algorithm Ester et al. (1996) on SMGI feature class to detect density clusters and add the cluster group in a new field of SMGI feature class		
Feature-based DBScan	Run DDBscan algorithm Ester et al. (1996) on SMGI feature class to detect density clusters for each group in SMGI feature class (e.g. User)		

(continued)

Table 1 (continued)

Spatial analysis of SMGI	
Tool name	Function
Advanced SQL maximum	SQL selection on a SMGI feature class to detect maximum values for a specific group (e.g. User cluster exposing max number of points)
Google™ static maps	Add the Google static map URL in a new field of SMGI feature class

picture of the overall perception of the users for each neighborhood at the local scale: the Castello neighborhood was depicted as the historical one, while the Marina as the retail and leisure one. As a matter of facts, both are historical neighborhoods, but the functional characterization clearly emerged in the analysis of the users perceptions. The results of the spatial analysis and the textual analysis are displayed in Fig. 2 and Table 2, respectively.

More interestingly, while it may be argued that this study on perception would be representative of the YouTube user community only, the same results were confirmed by another pilot study named Cagliari, I Care! 2.0, carried on using PIC! with a completely different pool of participants. In spite of different datasets, SMGI origin, involved users, and time period, the analyses gave the same results for the examined neighborhoods, raising interesting questions about SMGI reliability and community representativeness for analysis, which should be further investigated with new ad hoc studies.

In the last example of SMGI analytics with Spatext, the geography of Iglesias municipality (Italy) was investigated through the spatial and temporal distribution of Instagram photos. The purpose of the analysis was the identification and classification of SMGI clusters, relying on spatial and temporal component of users'

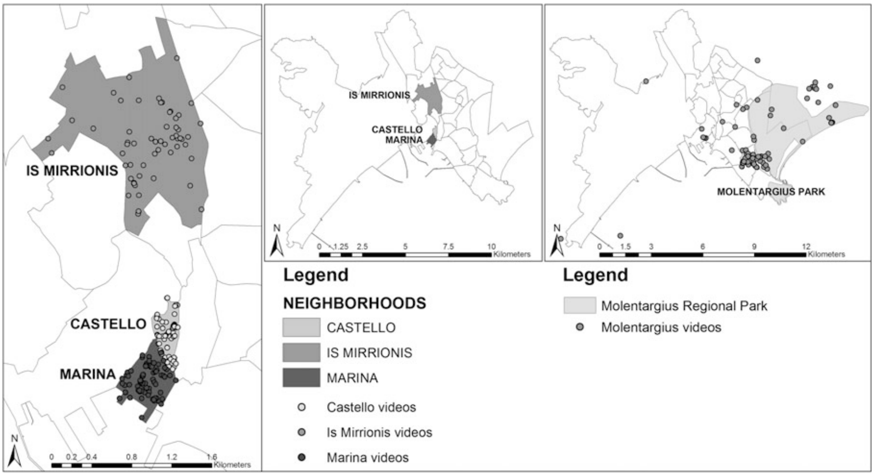

Fig. 2 Spatial distribution of videos related to Cagliari's neighborhoods

Table 2 Interpretation of major interests and concerns related to Cagliari's neighborhoods

Textual analysis of video title and description	
Neighborhoods	Keywords (frequency) translated
Castello	*Cagliari* (48), Urban (10), Safe (10), Concert (6), *Marmora* (6), Palace (5), Royal (5), Historical (4), Cathedral(2), Bastion (2)
Is Mirrionis	*Cagliari* (40), Tournament (10), *Monteclaro* (7), *CUS* (6), Soccer (4), Football Club (4), Sound (4), Final (4), Music (3), Park (2)
Marina	*Cagliari* (61), *Santa Lucia* (13), Concert (9), New Year's Day (6), Celebration (6), *San Sepolcro* (5), Music (5), Festival (4), Harbor (4), Church (4)
Molentargius natural park	*Cagliari* (62), *Quartu Sant'Elena* (15), Salt mine (10), Flamingo (8), Park (8), *Poetto* (5), Conference (5), Service (5), Pond (4), *Monte Urpinu* (3)

contributions, in order to detect buildings not appearing in official datasets. A one-year sample of approximately 14,000 photos from 1243 anonymized users was collected and georeferenced with Spatext, and analyzed in space and time. Spatext integrates the DBScan (Ester et al. 1996) and a Feature-based DBScan which were used to detect 290 clusters and 368 clusters, respectively. The difference between the results may be explained considering how the first tool relies exclusively on points' spatial density, while the second tool adds the grouping factor to the analysis, thus performing a separate DBScan analysis for points of each user in the SMGI dataset. Then, the Instagram dataset was integrated with the latest official buildings dataset available in the Regional Spatial DI. For each user's cluster the overlaps of the cluster centroid, of the photos barycenter, and of the shape with the buildings' footprint were evaluated. The analysis resulted in 113 clusters without any overlap, which were then visually assessed through the satellite photo viewer integrated in Spatext, in order to find un-mapped buildings in official information. An example of the results of the analysis is provided in Fig. 3, showing five different clusters (i.e. A, B, C, D and E), their centroids, the existing buildings footprints from the official dataset, and the Instagram point dataset. In the example, the manual investigation through the Google Map's satellite image allowed to detect two buildings which are not mapped in official dataset, identified by the cluster A and clusters C and D respectively. At the same time the analysis in the example confirmed the building in cluster B and allowed to detect a public space (cluster E), which is used by local residents for leisure. This example demonstrated how Instagram data may be used to elicit information related to topography of places and it may potentially integrate A-GI.

Altogether, these examples contribute to demonstrate how SMGI can be used to elicit information, not only about the physical geography of places to integrate existing A-GI, but also to express the perceptions of places and issues in time and spaces by the involved community, which may add a pluralist perspective of great relevance for spatial planning and decision-making.

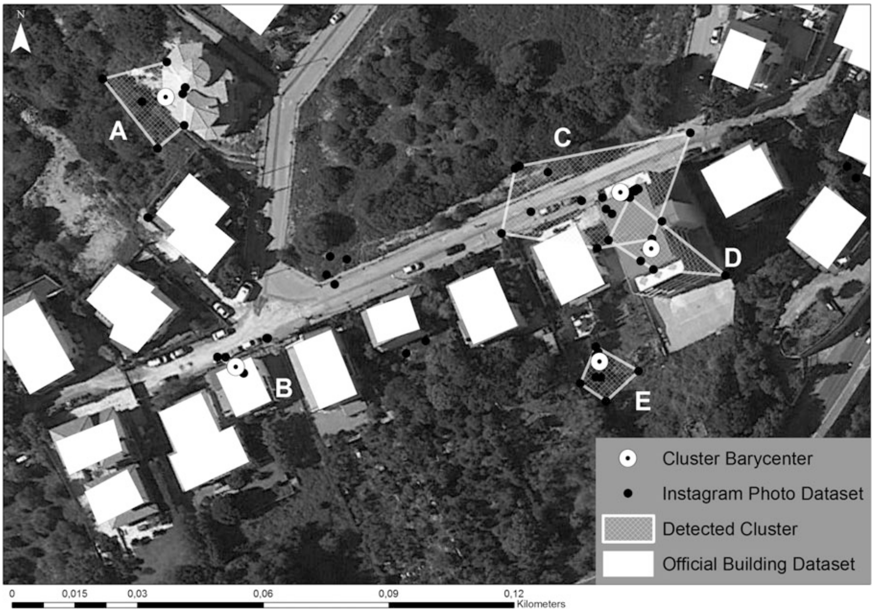

Fig. 3 Results of manual investigation on SMGI. Identification of buildings not mapped in the official dataset

3.3 The Use of SMGI in Tourism Planning

A wider case study was also developed in order to investigate the potential of the use of SMGI in tourism planning. The aim was to spatially analyze tourist preferences on (1) destinations and (2) tourism industry services, using tourists' ratings and comments collected by two major tourism social networks (i.e. Tripadvisor.com and Booking.com). Overall the study explores the following questions related to Sardinia's costumers' preferences:

- Which are the most popular destinations and why?
- What does attract tourists' attention and what do tourists appreciate/disregard?
- Why tourists choose those destinations?
- How this information can be used as tourism planning support?

From an operational perspective, the challenge was answering these questions relying on both data from the local SDI and from tourism SMGI, in order to discover novel and useful knowledge.

First of all, at the regional level, spatial analysis and statistics techniques for investigating the spatial distribution of tourists' preferences were carried on, identifying clusters of positive or negative preferences or hotspots of interest by tourist profiles. Then at the local level, large scale statistical and STTx analyses were carried on aiming at understanding, qualitatively, the reasons beneath patterns and singularities. Finally, a properly spatially calibrated model, analyzing spatial non-stationarity (Fotheringham et al. 2003), was implemented in order to express the impact of spatial variation in the relationships among dependent variable (tourists' preferences) and explanatory variables. Operationally, the study was carried on according to the following workflow:

- data collection and geocoding: data were extracted by TripAdvisor.com and Booking.com, geocoded, and integrated in a geodatabase including a one-year full set of data about 992 Tourism Lodging Services (TLS). The dataset includes TLS name, location, category, as well as quantitative and textual evaluations
- regional preferences dynamics analysis: spatial analysis and spatial statistics techniques, as well as STTx, were applied in order to detect and analyze preference clusters and hot/cold spots in Sardinia, using municipalities as unit of analysis
- local preferences dynamics analysis: data were further combined with spatial data themes from the regional SDI in order to earn deeper insights on the relationships among tourist preferences, local territorial features, and quality of industry services in selected destinations
- geographically weighted regression analysis: a model was developed in order to investigate how the detected patterns varied across different census tracts.

The last three steps are carried on iteratively on the relevant clusters and spots.

Spatial patterns of the TLS typology, together with their reviews analysis, offered interesting clues to characterize different destinations for tourism planning purposes. As an example, Cagliari TLS supply is characterized by the success of B&B, while for Alghero and Olbia, other two coastal city destinations, hotels and resorts are more popular among travelers. Analyzing tourists' nationality demonstrates also that tourists from Spain, Norway, and Ireland privileged the macro area of Alghero, while tourists from France and Germany preferred Cagliari, not surprisingly due to the majority of fly connections to and from these countries. In addition, Russian and Eastern European tourist showed a substantial preference for the Cagliari area.

Analyzing tourists' comments and evaluations enables investigation of the satisfaction level with destination and services. According to the results of the comments analysis, most words in the posts refer to spatial features and tourism structures, such as locations, hotel, or beaches. Frequent words also refer to the level of satisfaction with both destination leisure services and local traditions. Overall from the analysis, the main reason for tourists to visit Sardinia seems to be related to both its natural attractions and the presence of a unique cultural heritage.

The tourism preferences analyses in space were applied to investigate the preferences' patterns on both the territorial resources and the tourism industry features at the local level, across the whole region. The analysis started by mapping the Tourist Positive Preferences Incidence (TPPI) as the ratio between the positive scores and the number of TLS by municipality. The result shows an overall high and diffused concentration of preferences in the North-East coast. The Costa Smeralda area appears as the only area in Sardinia where tourist preferences overall fulfill visitors expectations. However, the analysis shows that the Alghero and the Cagliari areas also expose high TPPI rates.

After the spatial patterns analysis at the regional scale, the analyses shift to the local scale relying on spatial analysis and spatial statistics techniques on an integrated SMGI/A-GI geodatabase with the aim of investigating the local success factors within the single destinations. Alghero has been recognized as a best-selling destination from different tourists' typologies. Thus, the following question to answer was "What exactly does attract the tourists attention in Alghero and why?" Focusing on the local scale, the historic city center of Alghero clearly represents the major hotspot of tourists' attraction, while the modern residential districts in the outskirts represent a cold spot (Fig. 4).

The STTx analysis on tourists textual comments enables to understand both what tourists think and where. More than 880.000 long textual reviews were extracted (of which 1050 in English) related to Alghero. Table 3 indicates how the majority of the words in the posts refer to spatial or physical features, such as "location", "beaches" and "city center". Other frequent words are related to tourism structures, such as "hotel" and "staff". The textual analysis results also indicate a high satisfaction level with the destination leisure services. According to the results, the main reason for tourists to visit Alghero seems to be related to both its natural attractions, which include landscape features, such as beaches, and the presence of a unique cultural heritage. These facts generate a positive tourism destination image, which is the most influential psychological factor when tourists decide where to travel (Van Raaij 1986; Buhalis 2000).

Lastly, the study was also supported by the integration of SMGI with other A-GI describing topography, transport infrastructures, cultural heritage sites, and socio-economic features. The spatial relationships and the explanatory factors behind observed spatial patterns were modeled using the Geographic Weighted Regression analysis (GWR) (Fotheringham et al. ibidem , p. 9). The GWR is a technique which extends ordinary linear regression models by taking into account local variations in rates of changes.

In this case, the objective of the GWR analysis is twofold: the analysis is performed to investigate quantitatively why visitors' preferences are mostly located in the Alghero city center rather than in other locations in the municipality, and to discover which factors contribute to the Alghero high TPPI rate in those areas. The model was applied to a sample of 131 TLS distributed over 89 of the 471 census tracts in Alghero.

In this model, the dependent variable (local TPPI) was calculated as the sum of the positive score of the TLS normalized by the total number of comments per census tract. For each city census tract, the measure of each independent variable was calculated and normalized by the total area of the census tract. Preliminary results of the statistical tests suggested to exclude some of the explanatory variables originally chosen for the model, because they were not statistically significant. Eventually, the following candidate variables normalized by total area of census tract were included:

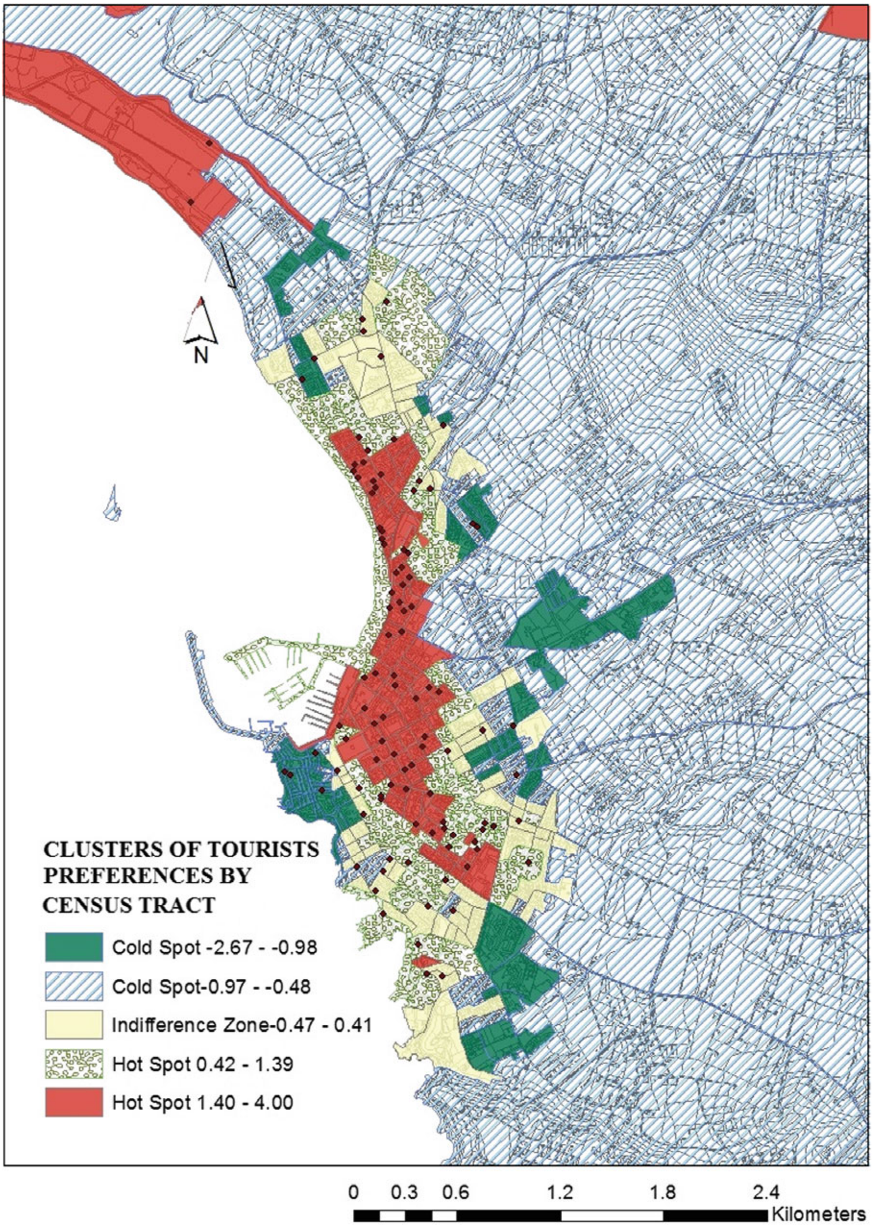

Fig. 4 Results of SMGI analysis: significant patterns of tourist positive preferences incidence in Alghero municipality by census tracts

Table 3 Results of the textual analysis: top 15 most used words in the Alghero cluster divided by category

Category	Words (frequency)
Geographic location	Location (1010), town (476)
Services	Staff (890), restaurant (643), room (459), hotel (469), pool (230), food (180)
Accessibility	Mapped (250), harbor (237), proximity (164), walking (146)
Natural and no natural components	City center (426), beach (378), old city (132)

1. number of historical buildings;
2. number of restaurants and facilities;
3. hectares of natural protected areas;
4. distance from the main transport nodes;
5. proximity to the historical city center;
6. distance from the most popular beach.

The assumption is that if the value of the TPPI at the local level is similar to the values that it takes in the closest spatial units, the variable is characterized by spatial autocorrelation. This issue can be addressed by adding a spatially-lagged dependent variable to the set of covariates (Anselin 1988, Anselin et al. 1996).

The presence of autocorrelation of the dependent variable (i.e. the normalized local TPPI) of a model is detected through the Moran's test. The local Moran's Index for the second order of queen contiguity was 0.06636, which is quite meaningful; the p-value was 0.0000146 ($0.01 \le p < 0.05$) hence statistically significant; the z-score featured positive sign, meaning that local spatial autocorrelation in the dependent variable is higher than it would occur randomly (Table 4). Thus, the regression is estimated using the six dependent variables and the tourists' preferences weight matrix obtained by the Moran's test.

In order to better understand the local variation of the explanatory variables, the coefficient raster surfaces created by GWR in ArcGIS were taken into account. The analysis allows to investigate how spatially consistent relationships between the dependent variable and each explanatory variable are across the study area. In addition, the analysis of the coefficient distribution as a surface shows where and how much variation exists. Figure 5 shows the residual values of the model, obtained by the differences between the fitted and the observed values of the dependent variable.

The map of the residual values is classified with the standard deviation method. As it is possible to note, for the majority of tracts, the values of the standardized

Table 4 Summary output of the Moran's index

R squared	Constant a	Std err a	t-stat a	p-value b	slope b	Std err b	t-stat b	p-value b
0.0393	0.0106	0.0151	0.703	0.482	0.0664	0.0151	4.38	0.0000146

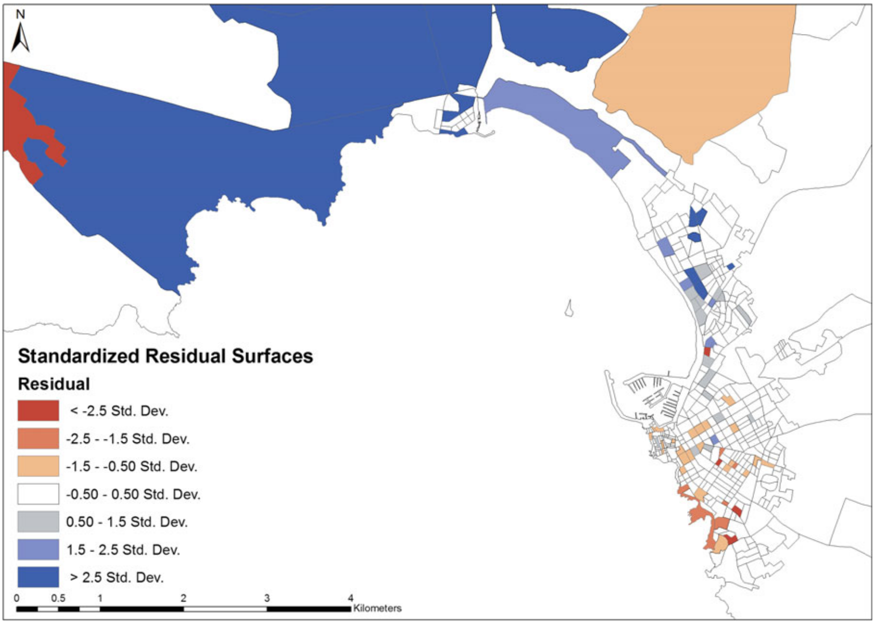

Fig. 5 GWR results: standardized residual surfaces

residuals are in a range between −1.5 and 1.5; few census parcels show standardized residual values higher than 1.5 or smaller than −1.5. Not surprisingly, many of the census tracts are placed in the city center, where the presence of historical buildings coupled with typical restaurants and leisure is more significant.

The results concerning the goodness of fit of the regression are significant: R-squared is very high, 0.943992, which indicates that the GWR model explains more than a 94 % of the variance of the positive tourists' preferences at the local level. The GWR model coefficients of the variables show the relationships between the dependent variable and each explanatory variable. The coefficients of the variables related to location are almost always significant (with p values less than 5 %) and show positive sign. The variables "hectares of natural protected areas" and "number of natural sites" are not significant, for the p value is higher than 10 %, while the variable "number of restaurants", related to service quality, shows a significant coefficient (with p value = 0.0167321) and positive sign. Table 5 shows the coefficient value for each considered variable.

Overall, these findings suggest that the spatial interest of the participant is quantitatively influenced by the chosen explanatory variables. The selected variables give a more or less significant contribution to the explanation of the tourists' preferences through their coefficient. In the model, the inclusion of the only tracts where the TLS are present, allows saying that the values of the coefficients in the tracts reflect the positive effects of the geographic position and the presence of facilities.

Table 5 Results of the GWR model: influence of each explanatory variable on dependent variable (normalized tourist's preference)

Variable	Coefficient	Std: Error	z-value	p-value
W_N_TPPI	0.058552	0.035614	1.644075	0.100161
CONSTANT	−0.012852	0.008821	−1.457048	0.145103
N_H_BUILD	0.753521	0.076918	9.796376	0.000000
DIST_BEACH	1.015170	0.100169	10.134590	0.000000
N_RESTAUR	0.089016	0.037206	2.392542	0.016732
H_NATURAL	0.005126	0.032138	0.159486	0.873286
D_AIRPORT	0.448991	0.126307	3.554754	0.000378
PROX_C_CENTRE	−0.005772	0.010428	−0.005536	0.005799

The results provide insights on the tourism preferences dynamics in Alghero, which would have been not possible to obtain through more traditional data sources for tourism planning. In addition, these findings confirm that the success of tourist destination is closely dependent not only on the quality of the tourist industry offer, but also on the territorial setting of the destinations, including the natural, cultural, and the physical character of the places, as well as the infrastructure and services. So far, the literature on TLS distribution dealt with several relevant sustainability issues, however often the spatial dimension of tourists' subjective perception was neglected. Including the latter dimension may open new opportunities for planners and offer new research challenges for a pluralist customer-oriented view on strategic tourism development.

4 Conclusion

The earlier results of this study offer an overview of possible uses of active and passive SMGI platforms to investigate what people observe, evaluate, and how they behave in space and time. The underlying endeavor of the study was to develop novel SMGI analytics methods and tools, which may help to access data sources and extract meaningful knowledge for spatial planning.

A number of case studies are discussed which demonstrate how SMGI, from active and passive sources, may be integrated with A-GI and used to understand people perceptions, contributing to define a pluralist model of local identity. In addition, several examples demonstrated how SMGI extracted from popular social media platforms may be used to detect changes in topography, as in the case of Instagram images, as well as, social and economic processes in the case of tourism planning. In the Sardinia tourism case study the underlying assumption was that the same methods and tools can be used successfully in urban and regional planning as much as in tourist planning, for in both cases they contribute to take into account a pluralist view on strategic development issues.

Overall the case studies show how qualitative and quantitative analysis can be applied to SMGI using spatial analysis and statistics combined with STTx techniques, which can be used to verify hypothesis unleashing the knowledge enclosed in the huge amount of qualitative descriptive social media comments.

Moreover, active and passive SMGI platforms may enable such scenarios where a city planner would aim to “listen” to what the local community feels about community issues or to interact with them to ask what alternative projects, or development options, would be welcome to the community. This would be a possible common use-case in the tradition of Public Participation/Participatory GIS (PPGIS) domain. However, until recently, PPGIS initiative required substantial endeavor in order to set-up technology and management. The availability of geographic social network platforms may ease the process both in technology and social terms. As a matter of facts, while on the one hand almost no technology set-up is needed, on the other hand the social network functions require less commitment by the potential participants, which increasingly use the media in their daily lives. The participation to the initiatives would blur with the everyday social networking activity. Accordingly, we can use such platforms as PIC! to create dynamic dialogues and monitor people interests about places on a routine base, or launch time-limited initiatives to ask the assessment of development alternatives.

A first conclusion may be that, as for the subject of observation (i.e. the features for places), the tools required to collect and extract knowledge are very local in nature. Globalization may occur in the future to this respect, but for the time being both A-GI and VGI sources are differently available in diverse places, contributing to define the local identity even in the knowledge communication media.

What appeared clear from the development, as well as from the usage of Spatext, is that any SMGI source is characterized by a specific data model and by different rates in geographic diffusion. Therefore different combinations of analytical approaches are required in order for the results to interpret the local context appropriately.

Several issues are still unresolved and need further investigation such as those related to data representativeness. Nevertheless, the presented case studies show how, in certain cases, similar images for a given place were depicted by different groups of users with different profiles, exposing invariance in the perception of places. While this issue should be further investigated, the potential of these new data sources already far exceeds those of more traditional community inquiring methods. To what extent SMGI may contribute to enhance the quality of knowledge, and eventually bring innovation to spatial planning, is still to be verified, but this question represents a very stimulating issue for further research.

Acknowledgments The authors wish to thank the anonymous reviewers and the editors for the valuable comments on the early version of this chapter. The work presented in this chapter was developed by the authors within the research project “Efficacia ed efficienza della governance paesaggistica e territoriale in Sardegna: il ruolo della VAS e delle IDT” [Efficacy and efficiency of landscape and environmental management in Sardinia: the role of SEA and of SDI] CUP: J81J11001420007 funded by the Autonomous Region of Sardinia under the Regional Law no. 7/2007 “Promozione della ricerca scientifica e dell’innovazione tecnologica in Sardegna”.

References

Anselin, L. (1988). *Spatial econometrics: Methods and models* (Vol. 4). Berlin: Springer.

Anselin, L., Bera, A. K., Florax, R., & Yoon, M. J. (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics, 26*(1), 77–104.

Ball, M. (2010). What's the distinction between crowdsourcing, volunteered geographic information and authoritative data? http://www.sensysmag.com/spatialsustain/whats-the-distinction-between-crowdsourcing-volunteered-geographic-information-and-authoritative-data.html. Accessed Jan 15, 2015.

Batty, M., Axhausen, K. W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., et al. (2012). Smart cities of the future. *European Physical Journal-Special Topics, 214*(1), 481–518.

Bruns, A. (2006). Towards produsage: Futures for user-led content production. In F. Sudweeks, H. Hrachovec & C. Ess, (Eds.), *Cultural attitudes towards communication and technology 2006* (pp. 275–284). June 28–July 1, Tartu, Estonia.

Buhalis, D. (2000). Marketing the competitive destination of the future. *Tourism Management, 21* (1), 97–116.

Campagna, M. (2014). The geographic turn in social media: opportunities for spatial planning and geodesign. In *Computational science and its applications—ICCSA 2014* (pp. 598–610). Berlin: Springer.

Caverlee, J. (2010). A few thoughts on the computational perspective. In *Presented during the specialist meeting on spatio-temporal constraints on social networks, December 2010*. Santa Barbara: CA, USA.

Craglia, M., Ostermann, F., & Spinsanti, L. (2012). Digital earth from vision to practice: Making sense of citizen-generated content. *International Journal of Digital Earth, 5*(5), 398–416.

Elwood, S., Goodchild, M. F., & Sui, D. Z. (2012). Researching volunteered geographic information: Spatial data, geographic research, and new social practice. *Annals of the Association of American Geographers, 102*(3), 571–590. doi:10.1080/00045608.2011.595657.

Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd* (Vol. 96, pp. 226–231).

Fotheringham, S. A., Brunsdon, C., & Charlton, M. (2003). *Geographically weighted regression: The analysis of spatially varying relationships*. New York: Wiley.

Goodchild, M. F. (2007). Citizens as voluntary sensors: Spatial data infrastructure in the world of web 2.0. *International Journal of Spatial Data Infrastructures Research, 2*, 24–32.

Goodchild, M. F., & Glennon, J. A. (2010). Crowdsourcing geographic information for disaster response: A research frontier. *International Journal of Digital Earth, 3*(3), 231–241.

Haklay, M. (2013). Citizen science and volunteered geographic information: Overview and typology of participation. In *Crowdsourcing geographic knowledge* (pp. 105–122). Netherlands: Springer.

Krumm, J., Davies, N., & Narayanaswami, C. (2008). User-generated content. *IEEE Pervasive Computing, 7*(4), 10–11.

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D., & Van Alstyne, M. (2009). Life in the network: The coming age of computational social science. *Science, 323* (5915), 721.

Nogueras-Iso, J., Zarazaga-Soria, F. J., Lacasta, J., Béjar, R., & Muro-Medrano, P. R. (2004). Metadata standard interoperability: Application in the geographic information domain. *Computers Environment and Urban Systems, 28*(6), 611–634.

Poser, K., & Dransch, D. (2010). Volunteered geographic information for disaster management with application to rapid flood damage estimation. *Geomatica, 64*(1), 89–98.

Roche, S., Nabian, N., Kloeckl, K., & Ratti, C. (2012). Are 'smart cities' smart enough? In *Global geospatial conference 2012*. German: Global Spatial Data Infrastructure Association.

Roche, S., Propeck-Zimmermann, E., & Mericskay, B. (2013). GeoWeb and crisis management: Issues and perspectives of volunteered geographic information. *GeoJournal, 78*(1), 21–40.

Schaffers, H., Guzmán, J. G., Navarro, M., & Merz, C. (2010). *Living labs for rural development: Results from the C@ R integrated project*. Madrid, Spain: TRAGSA.

Stefanidis, A., Crooks, A., & Radzikowski, J. (2013). Harvesting ambient geospatial information from social media feeds. *GeoJournal, 78*(2), 319–338.

Sui, D., Elwood, S., & Goodchild, M. (2013). Crowdsourcing geographic knowledge. Springer.

Toppeta, D. (2010). *The smart city vision: How innovation and ICT can build smart, "Liveable", sustainable cities, the innovation knowledge foundation*. http://www.inta-aivn.org/images/cc/Urbanism/background%20documents/Toppeta_Report_005_2010.pdf. Accessed Jan 21, 2015.

Van Raaij, W. (1986). Consumer research on tourism mental and behavioral constructs. *Annals of Tourism Research, 13*(1), 1–9.

Washburn, D., Sindhu, U. (2010). Helping CIOs understand "smart city" initiatives. Forrester Research, Inc. http://www.uwforum.org/upload/board/forrester_help_cios_smart_city.pdf. Accessed Apr 20,2015.

Zook, M., Graham, M., Shelton, T., & Gorman, S. (2010). Volunteered geographic information and crowdsourcing disaster relief: A case study of the Haitian earthquake. *World Medical and Health Policy, 2*(2), 7–33.