

CHAPTER 4

Social Media Geographic Information: Why social is special when it goes spatial?

Michele Campagna

Università di Cagliari, DICAAR, Via Marengo 2, Cagliari 09123, Italy,
campagna@unica.it

Abstract

This contribution introduces the concept of Social Media Geographic Information (SMGI) as a specific type of Volunteered Geographic Information (VGI). Unlike other kind of VGI, which may originate from geographic measurements crowdsourcing, SMGI brings in addition a special potential for it may express community perceptions, interests, needs, and behaviors. Hence, SMGI may represent an unprecedented resource for expressing pluralism in such domains as spatial planning, where it may convey the community collective preferences contributing to enrich knowledge able to inform design and decision making. In the light of these assumptions, the main issues relevant for SMGI collection and analytics are presented from the perspective of the spatial planning and governance domain, and a framework for the SMGI analytics in planning, design, and decision making is proposed.

Keywords

Social Media Geographic Information (SMGI), Spatial Data Infrastructure (SDI), Spatial Planning and Governance, Geodesign, Planning Intelligence.

How to cite this book chapter:

Campagna, M. 2016. Social Media Geographic Information: Why social is special when it goes spatial?. In: Capineri, C, Haklay, M, Huang, H, Antoniou, V, Kettunen, J, Ostermann, F and Purves, R. (eds.) *European Handbook of Crowdsourced Geographic Information*, Pp. 45–54. London: Ubiquity Press. DOI: <http://dx.doi.org/10.5334/bax.d>. License: CC-BY 4.0.

Introduction

It is often assumed, as a proxy benchmark for the growing amount of information being produced, that in 2009 more data were generated by individuals than in the entire history of mankind through to 2008. In this avalanche of information, social data are increasingly used to build balanced relationships between business and customers. This way, consumers are stimulated by successful web actors to share their data truthfully with the prize of earning the power of being listened to, in addition to other concrete advantages: the online world is beginning to be ruled by users' expectations (Weigend 2009) with major implication for industries and businesses. Likewise, social media are growingly becoming a major arena for politics. We live in a time where the premiere of a candidature to the United States of America presidential elections is expected to be released on Twitter and YouTube channels, and where social movements involving thousands of people are organized on the internet. More and more public authorities are moving online along the stream of a general digital social uptake. Nevertheless, unlike mainstream politics or other sectors of government, in the domain of spatial planning and governance methods and tools to fully exploit the potential of social data, actively (i.e. to create discourse) or passively (i.e. to listen to the community), still are not widely used. This poses questions concerning the actual willingness of citizens to establish such powerful user relationship with public authorities, and what public authorities can do to meet citizens' expectations.

From a technical perspective, nowadays many organizations rely on social media management tools to interact with their customers, which in the case of public authorities include the citizens; however these social engagements tools often remain siloed from other enterprise applications (Oracle 2013). As in many other domains in the private and the public sector, spatial planning and governance is a sub-domain where the integration of social data with authoritative official information may provide opportunities for improving the dialogue with the citizens, by not only listening to their preferences but also by monitoring the social processes they are involved in, towards more pluralist, informed, and community-oriented decision making.

Unstructured social media contents that capture citizens' interests, intentions, perceptions, and needs may enrich traditional institutional and other commercial data sources. Key Performance Indicators can be developed to monitor through real time dashboards the reactions of the community to the public policies and actions helping to respond promptly to citizens' behaviors, moving a step towards a new generation of planning intelligence, thereby contributing to a more sustainable and smart growth.

With this premise the chapter is articulated as follows. In the next section a brief overview of recent advances in digital spatial data sources argue knowledge building in planning is enriched by the availability of social data. Afterwards, a definition of Social Media Geographic Information (SMGI) is given as

a special type of Volunteered Geographic Information (VGI). Then, in the core section of the chapter, a novel SMGI analytics is proposed from the perspective of spatial planning and governance. The conclusions briefly summarize and propose issues for further research development.

Social data enrich the planning intelligence

Recent approaches to spatial planning and design propose the concept of Geodesign, emphasizing the role of knowledge about the local territorial context to inform design and decision-making. According to Steinitz (2012) there is no such profession as the Geodesigner, rather a Geodesign process is carried out through collaboration among different experts coming from the design disciplines and from the Geographic (Information) Sciences, as well as stakeholders and other actors from the local communities, or the people of the place. Unlike until a decade ago, such an approach is currently enabled thanks to development both in authoritative and volunteered sources of geographic information.

On the one hand, in an increasing number of countries and regions developments in Spatial Data Infrastructures (SDI) are starting to offer planners dozens, and in some cases even hundreds, of official large-scale spatial data layers, enabling the transition from analogue cartography analysis to digital geoprocessing in the representation and analysis of territorial processes as well as in the environmental impact assessment of design alternatives. This is the common case in Europe, where the Directive 02/2007/EU establishing the Infrastructure for SPatial InfoRmation in Europe (INSPIRE) is promoting public access to official spatial data produced by public authorities at all levels in the Member States. Along the adoption and the implementation of INSPIRE, in a growing number of regions, Advanced Regional SDIs are offering spatial data and services to professionals working on spatial planning and environmental impact assessment and are thus starting to bring innovation into the planning and design practice (Campagna & Craglia 2012).

On the other hand, the wealth of Volunteered Geographic Information (Goodchild 2007) offered by geobrowsers and widespread diffusion of GPS-equipped handheld devices, is starting to represent a novel –but already must-have – sources of information in many fields according to neo-geography or citizens science approaches. OpenStreetMap⁴ may be considered one of the most successful example of GI crowdsourcing to create a comprehensive and high quality open spatial dataset as a major alternative to more traditional official or commercial sources. Topography, networks, habitats, biodiversity, diseases spreading, climate change, and hazards are some of the examples of environmental and social processes being mapped by voluntary observers

⁴ www.openstreetmap.org.

acting as citizens sensors. However, only a fraction of available VGI is purposefully produced and contributed, while an even larger share is made available often as unaware results of the use of social media on web and mobile apps. With regard to the latter share, which is the focus of this paper, the next section argues that social media georeferenced content deserves to be treated individually in research for its peculiar characteristics, and special focus is given to its potential as complementary knowledge base in spatial planning and governance to support design and decision making.

Why social data are special when they are spatial?

Social Media Geographic Information (SMGI) can be defined as any piece or collection of multimedia data or information with explicit (i.e. coordinates) or implicit (i.e. place names or toponyms) geographic reference collected through the social networking web or mobile applications. Social data are acknowledged as a good of major value in the digital economy, and their potential for enhancing more traditional analytics is of the utmost importance. A big part of social data however also features spatial (and temporal) references, thus their integration with more traditional Authoritative Geographic Information (AGI) may enable a further step towards the next generation of geospatial intelligence.

SMGI is a sub-category of VGI and can be active or passive, depending on the type of application with which it is collected: applications purposefully created and/or used to collect SMGI in participatory initiatives (as in Campagna 2014) originate active SMGI, while SMGI harvested by general purpose social media such as Twitter or Instagram are passive, and can be considered more generically as user generated content.

Multimedia content of SMGI may include texts, images, videos, or audios in whatever combination usually aggregated in place marks or posts. Together with spatial references, place marks and posts usually feature a time reference and creator, or user owner.

Application Programming Interfaces (API) can be used by the public to access SMGI. Hence, the data model of each publicly accessible sources depends both on the original data model in the social media owner database and on the API. In general, the publicly available SMGI data model features a subset of the original attributes or multimedia data, implying that the analytical potential is in general greater within the social media companies than for the public (Lazer et al. 2009). Through the APIs, SMGI may be retrieved and accessed by keyword, by space, by time, by user, or by a combination of the former depending on the original social media platform and/or API. Data returned by a query through the API can be converted in a spatio-temporal dataset. Hence, spatial-temporal analyses are supported on SMGI as well as user-behavioral analysis (i.e. the analysis of user's behavior in term of data production and sharing). Spatio-temporal and user-behavioral analyses may be combined with querying and

mining techniques on multimedia (e.g. spatio-temporal textual analysis can be defined as the analysis of text in a given area in a given moment or time frame) creating new opportunities for SMGI analytics.

Towards SMGI Analytics for spatial planning and governance

As introduced in the previous sections, it is argued in this paper that SMGI may turn out to be a very valuable source of information to support spatial planning and design. However, a novel analytics is to be formalized for the peculiar data models which make this type of information different from more traditional vector spatial datasets, with which it can be integrated (i.e. AGI; e.g. spatial data layer from institutional SDI) in order to elicit knowledge useful for informing spatial planning, design, or governance.

At the current state of development, common AGI vector datasets available for download in national or regional SDIs as shapefiles or remotely accessible as Web Feature Services, in Italy as in other countries in Europe, feature a geographic and a thematic component or dimension. Hence, they can be represented as follows:

AGI = $\langle \mathbf{x}, \mathbf{y}, \mathbf{z}; \mathbf{a}_i \rangle$ where $\mathbf{x}, \mathbf{y}, \mathbf{z}$ represent the geographic coordinates, and \mathbf{a}_i any thematic alphanumeric attributes of the common relational database data types (i.e. text, incl. URL, numbers, or dates).

In contrast, SMGI usually features a richer data model including temporal and multimedia components. Additionally, each piece of information exhibits a user dimension, which may include an identifier as well as other data which convey information on the user's profile. The latter plays a special semantic role for the user who produced and shared the single piece of information becoming a dominant dimension from the analytical perspective. In addition, each piece of SMGI often features a score expressing agreement by or interest for and popularity within the virtual community, due to the functioning of the majority of the social networking apps.

Thus, SMGI can rather be represented as follows:

SMGI = $\langle \mathbf{x}, \mathbf{y}, \mathbf{z}; \mathbf{t}; \mathbf{u}; \mathbf{m}_i; \mathbf{l} \rangle$ where \mathbf{t} represents time associated to each element of the set, \mathbf{u} the user, \mathbf{m}_i the multimedia content (i.e. text, images, video, or audio clip), and \mathbf{l} the amount of 'likes and dislikes', the number of 'stars', or any other kind of popularity or agreement score, which indicates consensus on the measure and should be treated accordingly in the analysis. Figure 1 summarizes the main features of the AGI and SMGI different data models.

Therefore, any SMGI analytical framework should include not only traditional spatial analysis but also temporal, multimedia, and user behavioral analyses methods, and these should be tightly integrated in order to fully exploit the knowledge potential embedded in data. From a planning analysis perspective, coupling these methods in an integrated GIS application would be an advantage in as much GIS is (becoming) the common platform for the

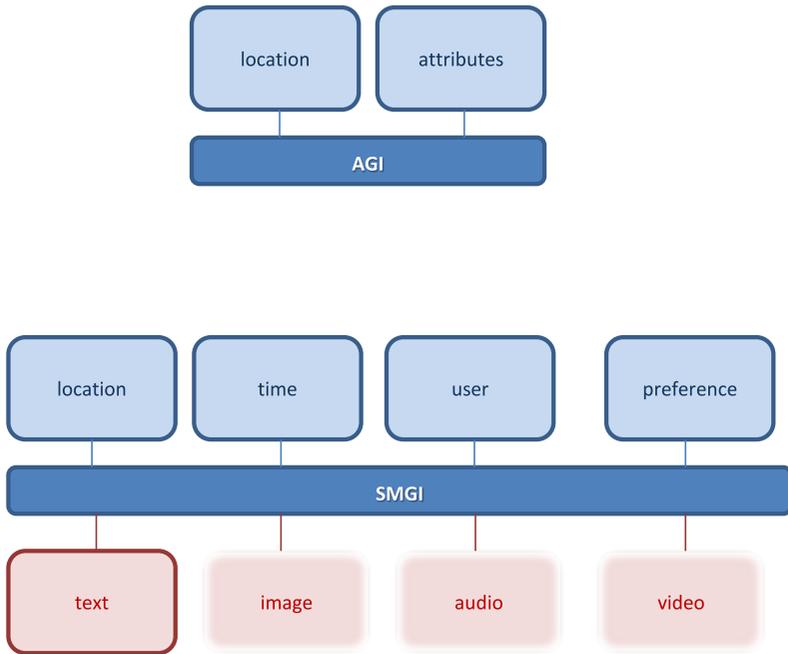


Figure 1: Scheme of the AGI (up) and of the SMGI (down) data models.

planning profession given the role of maps in expressing knowledge and design in this domain. With this consideration in mind, a framework for SMGI analytics have been developed by the author with the objective of exploiting this new resource of information in order to enrich the knowledge about the local context from the social perspective and to better support spatial planning and governance. This framework is under development; nevertheless the current results seem promising and allow proposing a tentative formalization as a guide for further research in this field.

In the remainder of the paper, this tentative framework for SMGI analytics is presented with references to the case studies developed at the University of Cagliari under the supervision of the author. The case studies relate to two major research streams. The first one pertains the development and use of a map-based social networking platform, namely 'Place, I care!' (PIC! 1.0). The latter can be defined as an active SMGI resource where users can create thanks to a user-friendly interface a private or public project. Thanks to an advanced user permissions manager each project can be customized to the contextual use case requirements. This way, accepted users can be allowed (or not) either to post, like/dislike, or comment, enabling different levels of participation in a map based discussion. Simple query functions are available in the map interface and collected data can be easily exported for further analysis with GIS

packages. PIC! 1.0 was used in a number of pilot projects (Campagna 2014; Campagna et al. 2015).

The second research stream concerns the collection and use of SMGI produced by major social networking platform such as Twitter, YouTube, Instagram, and Booking.com (i.e. passive SMGI resources). In both cases, SMGI datasets were collected in selected areas at various scales (i.e. from the global, to the regional, to the local) and then integrated with other sources of AGI (e.g. from regional or local SDI) or VGI (e.g. WikiMapia).

When SMGI is integrated in a GIS project with other authoritative and volunteered spatial datasets, the peculiarities of their data models enable the analyst to perform the following analysis on SMGI data:

- **Spatial analysis of user interests:** thanks to the widespread use of social media, the high number of georeferenced posts enables us to investigate the patterns of user interest in space by density (Campagna 2014) and clustering functions (Massa & Campagna 2015). The overlay with topographic AGI such as administrative boundaries, or physical artefacts such as buildings, infrastructure, services or public spaces, may offer useful hints to public authorities to understand not only which places are important to the community and how they are perceived (Campagna 2014), and by whom the community is eventually composed (e.g. local people, commuters, tourist or other);
- **Temporal analysis of user interests:** the temporal reference is often an available attribute in SMGI, enabling to study when given regional destinations, urban districts, public spaces, or other infrastructures and services are used along the year, the months, the week, or the day (an example of this type analysis is given in Massa and Campagna, in this volume);
- **Spatial Statistics of user preferences:** collecting posts by spatial units enables planners to analyze patterns in user interest at different scales. An example is given in Floris and Campagna (2014), where the hot-spot analysis has been used at the regional level to study the distribution by municipality of positive user assessments by user profiles, where the hot-spot analysis has been used at the regional level to study the distribution of positive user assessments by user profiles to discover where young vs. elder, or family vs. solo tourists prefer to go during their trips. In the same study, the hot-spots were then investigated with Spatio-Temporal Textual analysis (see STTx below) and with geographically weighted regression analysis to explore at the local level what physical and locational factors may affect the preferences;
- **Multimedia content analysis on texts, images, video, or audio:** multimedia analysis is well developed in the case of texts analytics. However, it is currently more difficult to automatically extract useful information from images, video, or audio. In the case of text, many software packages can be used to apply simple (i.e. calculating words frequency, or tag clouds) to more advanced (e.g. sentiment analysis) text analysis techniques. These

techniques can be easily applied to subsets of SMGI obtained by spatial, temporal, or user query (see STTx below);

- **User behavioral analysis:** querying SMGI by user enables to study users' behavior in space and time. This information can be used to analyze, for example, whether a public space is visited by local people or by visitors. This information may be useful also for profiling: for the users visiting a certain place or service user Spatio-temporal footprints can be defined to identify people who mainly move locally, regionally, or internationally, and where they come from;

or

- **A combination of two or more of the previous** such as in **Spatial-Temporal Textual Analysis (STTx)**: textual analysis functions can be integrated within GIS applications (Campagna 2014), enabling the application of text analysis techniques to subset of SMGI selected by space and time. Several examples were tested by the authors to analysis the perception of different neighborhoods in the city using Place, I care! data (Campagna 2014), the judgment of tourists on a destination using Booking.com data (Floris & Campagna, 2014), or who is talking about an event of global reach around the world using Twitter and YouTube data (Massa & Campagna 2014). Although not verified yet by a systematic analysis, several case studies on the application of STTx to the same areas with different SMGI sources, thus different type of users, returned similar results, suggesting further research should be devoted to better understand the issue of representativeness.

Conclusions

The tentative framework presented in this paper derives from testing single SMGI sources integrated with AGI in a GIS environment. The research may be considered still in its infancy and the SMGI analytics framework proposed here is likely to evolve substantially in the future, nevertheless the potential seems already be very promising and it offers many issues to be further investigated. Early experiments seem already to demonstrate how it is possible to introduce new dimensions of analysis in order to build useful knowledge for design and decision making in spatial planning and governance. Further research is under development in order to explore the potential of integrating multiple SMGI data sources together. Indeed, each SMGI source has different peculiarities given by specific data models and public accessibility feature. In addition, each SMGI source has different rate of diffusion and usage in different regions making their combined use unique to a given place, making the case for local SMGI–AGI mixes. This issue is further amplified by the fact that also AGI sources may vary differently in different regions and countries. Nonetheless, the possibility of introducing pluralist knowledge on people's perceptions, preferences, or needs is an opportunity of utmost importance to bring innovation

towards more sustainable, democratic, and community-oriented design and decision-making in spatial planning and governance.

The case studies referenced have already demonstrated how SMGI data enable us to observe to the community and to establish a dialogue. From the planning perspective, the possibility to tightly couple participatory initiatives to traditional planning knowledge in an integrate environment is a promising frontier to be further investigated. Territorial marketing, urban, and regional as well as sector planning and spatial governance in general have now a new power to inquiry the limit of which seems to be only the ability to look at data from a new perspective and ask smart questions.

Acknowledgments

The results on which part of this contribution is based were developed 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 n° 7/2007 'Promozione della ricerca scientifica e dell'innovazione tecnologica in Sardegna'. The author wishes to thank a number of young researchers for their contribution to this project including Pierangelo Massa, Roberta Floris, Anastacia Girsheva, Konstantin Ivanov, Roberta Falqui e Sara Mura.

References

- Campagna, M. 2014. The geographic turn in Social Media: opportunities for spatial planning and Geodesign. In: Murgante, B., et al. (Eds.) *ICCSA 2014, Part II, Lectures Notes in Computer Science* 8580. Springer International Publishing Switzerland: pp. 598610.
- Campagna, M., & Craglia, M. 2012. The socioeconomic impact of the spatial data infrastructure of Lombardy *Environment and Planning B: Planning and Design*, 39(6): 1069–1083.
- Campagna, M., Floris, R., Massa, P., Girsheva, A., & Ivanov, K. 2015. The Role of Social Media Geographic Information (SMGI) in Spatial Planning. In: Geertman, S., Ferreira, Jr., J., Goodspeed, R., & Stillwell, J. (Eds.) *Planning Support Systems and Smart Cities*. Springer, ISBN 978-3-319-18367-1.
- Floris, R., & Campagna, M. 2014. Tourism Planning in Social Media Data: Analysing Tourists Satisfaction in Space and Time *TeMA – Journal of Land Use, Mobility and Environment* Special Issue Eighth International Conference INPUT Smart City Planning for Energy, Transportation and Sustainability of the Urban System, Naples, 46 (June 2014).

- Goodchild, M. 2007. Citizens as sensors: the world of volunteered geography *GeoJournal*, 69(4): 211–221.
- Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabasi, A.L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., & Van Alstyne, M. 2009. Computational Social Science. *Science*, 323: 721–723.
- Massa, P., & Campagna, M. 2014. Social Media Geographic Information: Current developments and opportunities in urban and regional planning. In: *Proceedings of the 19th International Conference on Urban Planning and Regional Development in the Information Society GeoMultimedia 2014*.
- Massa, P., & Campagna, M. 2014. Social Media Geographic Information: recent findings and opportunities for smart spatial planning *TeMA – Journal of Land Use, Mobility and Environment* Special Issue Eighth International Conference INPUT Smart City Planning for Energy, Transportation and Sustainability of the Urban System, Naples, 46 (June 2014).
- Massa, P., & Campagna, M. 2015. Integrating Authoritative and Volunteered Geographic Information for spatial planning.. In: Capineri, et al. (Eds.) *European Handbook of crowdsourced information*. London: Ubiquity Press, pp. XX–XX.
- ORACLE. 2013. The value of social data *ORACLE White paper*. Available at: <https://go.oracle.com/LP=1320> (Last accessed 28 June 2015).
- Steinitz, C. 2012. A Framework for Geodesign. Redlands, CA: Esri Press.
- Weigend, A. 2009. The social data revolutions. Harvard Business Review. Available at: <https://hbr.org/2009/05/the-social-data-revolution> (Last accessed 28 June 2015).