Chapter 2
We Live in Informational Landscapes

Abstract In this chapter, we reflect upon how the urban experience is nowadays affected by data produced through interconnected, digital ecosystems such as the user-generated content on social media. We will present the scientific literature related to studies, research projects, and commercial platforms that use geo-located social mining techniques to extract urban knowledge. We will also analyze the state of the art in this thematic field and identify opportunities for further investigation.

2.1 An Introduction to the Data City

A massive amount of information is emerging from the digitization of contemporary cities, providing novel means of understanding our lives and the inner dynamics of organizations and societies. City users can now provide local knowledge through a bottom-up approach, contributing information through the use of widely diffused technologies such as tablets, mobile phones, and similar devices. By posting pictures, checking in at specific places, and sharing geo-referenced content as part of their everyday experiences, people are—more and more—making information available on how they live in, use, and perceive urban spaces.

Over the past few decades, several scholars have offered descriptions that combine traditional city representations with new informational membranes hovering above urban fabrics. Since the early 1990s, sociologists and economists have been looking at contemporary cities as expressions of networks and fluxes of information (Castells 1989). Concepts such as "mediascape" (Appadurai 1990) and "informational landscape" (Graham 1998; Zook 2000) are essential in contemporary debates about cities, offering visions of places shaped and defined in terms of perspectival constructs, where digital data and information coming from multiple sources (sensing technologies, ubiquitous computing, and user-generated content) play an important role in shaping urban experiences.

P. Ciuccarelli et al., *Visualizing the Data City*, PoliMI SpringerBriefs,
DOI: 10.1007/978-3-319-02195-9_2, © The Author(s) 2014

The vision of a soft city based on concepts such as emergence, geo-bodies, and intelligent machines and where physical and digital dimensions are strictly interrelated becomes an important component of contemporary urban geography (Pickles 2004).

The underlying idea of these approaches is to view the urban experience as tied to the multiple, fragmented, and temporary layers of data and information generated by human-place interactions. This is what we define as the *data city*.

These data can be produced either collectively or by the individual; they can be aggregated or discrete, open or protected. They constitute observation points that allow the interpretation and description of behavioral patterns within specific temporal and spatial coordinates.

In this context, the challenge for urban studies is the integration of available digital databases with traditional data, aimed at capturing the variety of changes in urban practices. As Zook and Graham (2007) noticed, traditional methods for registering users' perceptions and activities in cities—such as surveys and ethnographic reports—are becoming inadequate to meet our contemporary society's need for information because they require a considerable amount of resources, in terms of time and money, and because they are often only tied to a specific and limited time (the period of investigation).

The recent technological developments in terms of ubiquitous connectivity and the emergent participation of Internet users in terms of social interaction are leading toward a redefinition of the possibilities of gathering and sharing firsthand information that can be analyzed to complement more traditional urban observations and surveys. These ephemeral and overlapping layers of information demand new modes of inquiry and synthesis.

A new generation of dynamic city representations capable of defining and visualizing both physical and social environments, as well as individual and collective narratives, is required. This stratification of experiences leads to new kinds of geographies, partially handled by citizens themselves: a diverse sets of practices that operate outside or alongside professional geographers (Szott 2006).

Depicting the data city means to present urban spaces through visual approaches that are able to capture their flows and bring them back in the form of static or dynamic images. This can offer new perspectives in the way city services, processes, and strategies are designed and implemented.

2.2 People-Generated Landscapes: Potentialities of Social Media Data

Among all the sources of information available at the urban scale, we are particularly interested in user-generated content.

The OECD (2007) defines content as "user-generated" when it is produced by non-professional persons (i.e., amateurs), as opposed to professional media producers. An interesting subset of user-generated content is the information

specifically produced and shared through social media services. The term *social media* refers to the means of interaction among people in which information is created, shared, and exchanged, almost often in virtual communities and digital networks (Ahlqvist et al. 2008; Morgan et al. 2012). Examples of social media are platforms such as Twitter, Facebook, Flickr, Foursquare, and Instagram.

At present, social media is still growing rapidly and—as highlighted by the 2012 Nielsen Report on social media (Nielsen 2012)—it has become an integral part of our daily lives, making social networking a truly global phenomenon. Social media has been described as a ubiquitous tool for social interaction. We would be wrong in assuming that it is only an activity for the younger generations. Although the penetration among 16–24 year olds remains the highest, it is the 25-to 45-year-old sector that has seen the greatest increase in usage in recent years (WAVE6 research 2012).

If on one side there is a growing concern about privacy, on the other side, an increasing number of people say that social networks are a crucial component of both their private life and their professional activities, and as such, the benefits of their use overcome any worry about privacy. A significant number of citizens, especially in North America and Europe, are sharing more data than ever before through photos, videos, or simple status updates on their profiles on social media. Moreover, recent data from Qualcomm (2013) suggest that emerging regions—including Latin America, China, and India—are witnessing substantially higher numbers of new smartphone subscriptions than North America, Japan, Korea, and Europe.

Even if the personal computer is still at the center of the social networking experience, people are increasingly looking to other devices to connect to social media: the average time spent on mobile apps and the mobile web accounts for 63 % of the year-on-year growth in overall time spent using social media. In 2012, 46 % of the users accessed social media through their smartphone; 16 % used a tablet. With more connectivity, people have more opportunities to use social media wherever and whenever they want (Nielsen 2012).

This layer of digital information that we constantly produce can be increasingly related to specific geographic locations through the global positioning systems (GPS) integrated in current mobile navigation devices and through the geographic markers that some social media leave on user-generated content. The combination of these factors produces and disseminates an immense amount of geographical information, which can be contributed voluntarily or involuntarily.

Physical spaces become places as they are experienced, lived, shared, and communally interpreted. These places become "practiced places," to quote de Certeau, or in other words, they become spaces interpreted and lived through the experiences of situated subjects (de Certeau 1984). These experiences are to some extent captured, shared, and communally interpreted through these social media applications that are capable of recording and geo-referencing a variety of signals. This transforms human beings into potential "sensors" that not only have the ability to process and interpret what they feel and think but also to geographically localize the information (sometimes involuntarily) and spread it globally through the Internet, thus drawing people-generated landscapes.

Our assumption is that by conducting an analysis of these people-generated landscapes—more specifically, of data sets based on information extracted from social media—we can recognize multiple interpretations of the city as they emerge, overlap, and influence each other and as they unfold from users' mental representations and spatial experiences of city spaces.

Through our research, we are mainly investigating the use of the types of social media data that are most likely to provide precise geo-localized information, such as Twitter, Foursquare, and Instagram. Although there are many social media platforms that provide geo-localized information, Twitter, Foursquare, and Instagram proved to be the richest sources in terms of the number of contributions for the geographic locations investigated in our two projects, Telltale and Urban Sensing.

Twitter is a micro-blogging service allowing its users to publish and share short texts (maximum 140 characters), generating online conversations with other users. Twitter contributions shared through mobile devices have a precise geo-localization and timestamp. Although Twitter's main focus is not geo-localizing activities, it is possible to perform an analysis to identity potential correlations between the types of issues that people talk about (content analysis) and the places where contributions originate. In March 2013, Twitter counted over 200 million active users creating over 400 million tweets each day (Twitter official blog 2013).

The main use of Foursquare is to share and show real-time information on the geographic positioning of its users. Places such as clubs, stores, public offices, monuments, and the like (called "venues") are categorized and shared by the users. Users can "check-in" at a given location just by selecting it from a list of venues their mobile application locates nearby, or by creating new venues not already present in Foursquare (e.g., a new bar or a new store). Venues are categorized by the main types of activities associated with them (categories such as: art and entertainment, food, shopping, residences, professional venues, and other places), and each user can also insert comments and pictures of the venue. As of September 2013, the company had 40 million users and a database of 55 million locations throughout the world (Cnet News 2013). Male and female users are equally represented, and 50 % of users are outside the US (Wikipedia).

Instagram is a social network based on photo sharing. The application enables users to take a photo, apply digital filters to modify it, tag the picture, and then share it with other Instagram users. Each picture has a precise geo-localization and timestamp, as do all the other contributions such as users' comments on pictures. As of September 2012, Instagram had 100 million registered users (Wikipedia).

2.3 Applications Targeting Urban Issues

Decision-making processes at an urban scale are normally fed by two types of knowledge: institutional knowledge, that is, knowledge possessed by urban institutions in different forms (documents, statistical data, plans, reports, speeches,

etc.), and local knowledge held by different groups of citizens who look at the urban environment through their daily experience, filtered by their personal competences and skills (Healey 2007). The interaction between these two sources of knowledge makes the decision-making system work. The interplay between institutional and local knowledge also depends on the specific configurations of power and authority that allow (or do not allow) dialogic processes between the different kinds of knowledge.

While institutional knowledge is identifiable, since it is reported, documented, and stored within each institution, local knowledge is distributed across different individuals and social groups, and for this reason, it can be less easily retrieved and operationalized.

In the last decade, the need for an increasing public participation has made its way into planning initiatives—such as open meetings, co-design sessions, and digital collaborative platforms—in order to gather this local knowledge (Hanzi 2007; Simonsen and Robertson 2013). Decision-making processes tend to be more open to the public due to the pressure exercised by the social demand of local communities to participate in public decisions and due to the fact that localized knowledge is essential in solving problems in our complex urban society, where the top-down recipes do not work anymore (Innes and Booher 2010).

Dealing with larger publics—even thousands of diverse people—and gathering their local knowledge requires new tools and techniques.

Emerging cartographic practices have proposed new models to represent and describe the city, and they stress the collaborative dynamics of the processes (Crang 1996; Schein 1997; Cosgrove 1999; Duncan and Duncan 2003; among others). A set of participatory platforms allowing individuals to share their experiences and ideas about public spaces is emerging; some examples are projects such as: Neighborhood information System, Boston Citizens Connect, Crimebase Philadelphia's Neighborhood information System, Urban API, Hubway (a bike-sharing system in Boston), Citysourced, UrbanSim, London City Dashboard, Live Singapore, Public Information Map, and City Tracking (links to the websites of these projects are listed in the references section).

These and similar experiences focus on specific areas of urban interest that can potentially be investigated through geo-located social media analysis. We present below some of the areas commonly explored by these kinds of platforms and some of the questions they target.

2.3.1 Characterizations of Geographic Areas

What is that makes people talk about an area? Is it theatres? Is it the food scene?

What are the keywords or sets of keywords that recur more frequently when users tweet from specific locations? Can these sets of keywords offer a characterization of these geographic areas?

Through the application of sentiment analysis techniques (Liu 2010), it is also possible to identify the emotional attitudes of the citizens toward their cities, specific neighborhoods within the cities, or single urban infrastructures and urban projects. Sentiment analysis allows researchers to extract information from a text (e.g., a tweet) and to detect users' emotional status, classified according to basic polarity (positive, negative, or neutral).

2.3.2 Temporal Dimension

How do users' contributions on social media change over time (on a daily, weekly, monthly, and yearly basis)?

Contemporary urban analyses are increasingly focused on micro and interstitial investigations to explore the temporary, precarious, and conflicting meanings attributed by specific groups to places with which we often maintain a familiar relationship: a neighborhood in our city, a park, an event, or a series of paths. Bringing a temporal dimension into those data gives us the possibility to compare their distribution and their peaks over time and to identify patterns.

The temporal dimension also allows us to trace users' movements over time, thus detecting different patterns for different categories of users such as tourists or city inhabitants. Where do people come from? How do different categories of users distribute themselves within the city in normal days or during a special event?

2.3.3 Places Identities and Relationships Among Geographic Areas

Which are the most mentioned city places in social media? Which places are never mentioned? Why? Where and when do people recognize a territory as peripheral? Which are the boundaries of the actual "center" of the city, as named by people? How can we define and how do people name areas in transition?

Moreover, analyzing the profile of users that contribute on social media is a promising method to intercept and predict characteristics of urban places, such as gentrification areas, as well as emergent places.

Furthermore, how can we intercept emergent links between different areas of the city? Which are the patterns of mobility of certain groups of citizens that inhabit or cross specific places?

2.3.4 Emerging Behavior

Very often people are not offered what they need from public institutions and administrations. This leads to emergent bottom-up initiatives meant to answer

eluded questions, needs, and desires. Can these bottom-up initiatives be identified through social media analysis?

Is there any conversation between these initiatives and the local structures of government? Is it possible to discover the lack of infrastructure and services of a city by analyzing users' online complaints? In Milan, there are no buildings officially used as mosques. Where do people talk then about religion?

2.3.5 Political Attitudes: Acceptance/Feeling Toward Local Policies and Urban Interventions

How does a population (social network users) react to new policies? The concept of customer satisfaction is often used to categorize the ways in which these reactions are expressed. Two primary indicators can be applied: one determining the "attitude" (e.g., Is the general feeling toward the introduction of a new pedestrian area positive or negative? Are people happy with the new bike lanes?) and one that attempts to gauge a more interpretative content of text on a discursive level (e.g., How are people talking about a new proposed policy intervention? Can we distinguish differences among different ethnic groups, social groups, profiles and gender, ages, or from people who inhabit different parts of the city?).

2.4 Existing Work

Several projects have tried to reflect on the potential of these new data sources. Cartographic projects based on different kinds of geo-located user-generated content and social media data have been conducted both by research institutions (e.g., *CASA* at the University College London, *SIDL Lab* at the Columbia University, *Senseable City Lab* at the Massachusetts Institute of Technology, and *Urban Age* at the London School of Economics) and independent scholars and design firms (e.g., Christian Nold or *Stamen Design*).

We cannot here present a complete and exhaustive review of the many important projects in the field. However, we highlight below some experiences that we consider relevant for the purposes of our research. The projects show how different social media streams can be analyzed using different strategies, both in terms of indicators and solutions to visualize them.

- *World's Eye* (2008) is one of the most pioneering projects related to our research. It was organized by the MIT Senseable City Lab at the Massachusetts Institute of Technology (Fig. 2.1). The project retrieved vast numbers of photos taken by thousands of users in the most photographed cities in the world (Barcelona, New York, and Rome) and uploaded the photos to Flickr. Using timestamps, coordinates, and tags provided by users, the Senseable City Lab

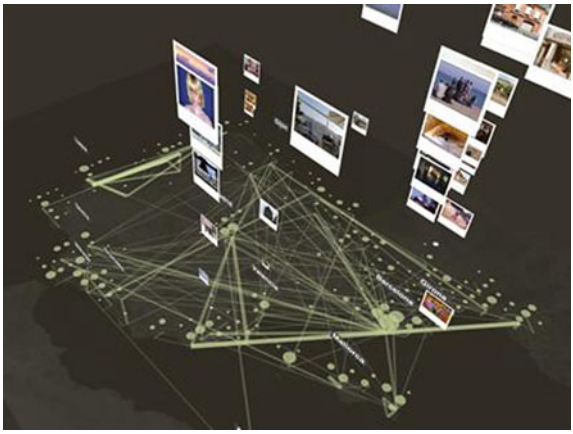

Fig. 2.1 From *World's Eye*, 2008, by MIT Senseable City Lab; Carlo Ratti, Assaf Bidermann, Fabien Girardin, David Lu, Andrea Vaccari, and Ernesto Arroyo (reproduced with permission)

designed geo-visualizations that reveal the movements of people through the points of interest of the three cities, including monuments, museums, theatres, and the like.

The project also produced several indicators that could be used to compare the evolution of the attractiveness of different city areas.

- *Connecticity* is an artistic and research project initiated at the end of 2008 by Salvatore Iaconesi and Oriana Persico. Connecticity aims at understanding the transformation of cities and urban spaces with the advent of ubiquitous technologies and networks. Connecticity mostly unfolded through a series of design experiments, some of which presented digital platforms that gather and analyze data coming from social networks and plot them on interactive maps.

For example, in the *VersuS* experiment—conducted in Turin (Italy) in 2011 within the Connecticity project—social media data streams were analyzed in order to identify the language the users used while contributing their user-generated content (UGC) from specific areas of Turin. The platform then created interactive maps that highlighted the geographic areas where a significant number of contributions using the same language emerged (Fig. 2.2).

- The *Geography of Buzz*, developed by the Spatial Information Design Lab at Columbia University, mapped the cultural epicenters of Los Angeles and New York by surveying 300,000 snapshots taken from the Getty Images archives relating to cultural events such as parties, openings, and premieres (Fig. 2.3). For example, their analysis showed that the "buzziest" neighborhoods in New York (i.e., areas presenting the highest number of contributions on social media according to a specific topic) were central and populated areas such as Times Square and SoHo and not the blooming artsy and newly gentrified areas. Williams, one the authors of the project, stated in a *New York Times* interview (2009): "We're going to see more research that's using these types of finer-grained data sets, what I call data shadows, the traces that we leave behind as we go through the city. They're going to be important in uncovering what makes

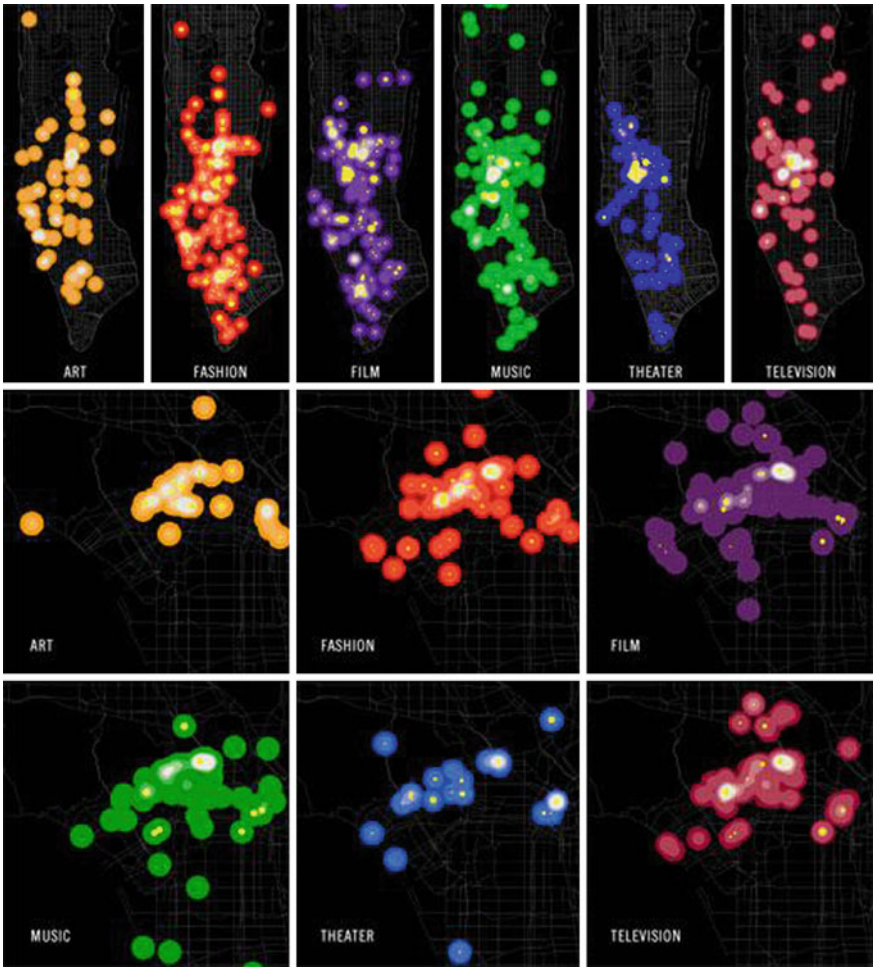

Fig. 2.2 *The Geography of Buzz*: *Art, culture, and the social milieu in Los Angeles and New York*, co-authored by Elizabeth Currid and Sarah Williams for the *Journal of Economic Geography* (2009) (reproduced with permission)

cities so dynamic." Currid, another author of the project, added in the same interview: "People talk about the end of place and how everything is really digital. In fact, buzz is created in places, and this data tells us how this happens."

- *Livehoods* (2012) is a recent research project conducted by the School of Computer Science at Carnegie Mellon University. Livehoods traces new geographic areas in the city by analyzing the check-ins on Foursquare. The technological system behind the Livehoods project dynamically creates geographical clusters and draws the consequent areas' borders originating from similar patterns in these check-in activities.

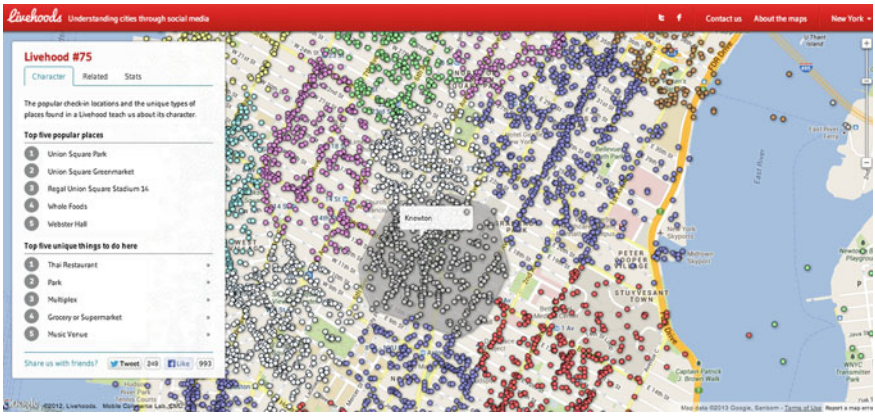

Fig. 2.3 From *Livehoods* by Justin Cranshaw, Raz Schwartz, Jason I. Hong, and Norman Sadeh (2012) (reproduced with permission)

- *Phototrails* (2013) is the result of collaboration between the department of History of Art and Architecture at the University of Pittsburgh, the Software Studies Initiative at the California Institute for Telecommunication and Information, and The Graduate Center, City University of New York (Fig. 2.4).

Phototrails analyzes user-generated data (such as Instagram pictures) using algorithms to detect visual patterns, dynamics, and similar structures. Using a sample of 2.3 million Instagram photos from 13 cities around the world, Phototrails shows how temporal changes in the number of shared photos, their locations, and visual characteristics can uncover social, cultural, and political insights about people's activity.

2.5 Specificities of the Research Trajectory

Our research trajectory departs from the experiences presented above.

The existing work constitutes an important reflection about the potential of georeferenced social media data as a means of investigating urban dynamics. In addition, some of the projects presented above also suggest the need to have a robust technological architecture in order to process a relevant quantity of data of this kind.

Thus, the original contribution of our research trajectory is articulated across the following directions:

- We aim at providing a wide range of stakeholders with a newly designed and implemented web-based and easy-to-use sensing platform that offers dynamic visual information. The platform offers fully customizable visual representations, made available as services. This will allow for outputs to show

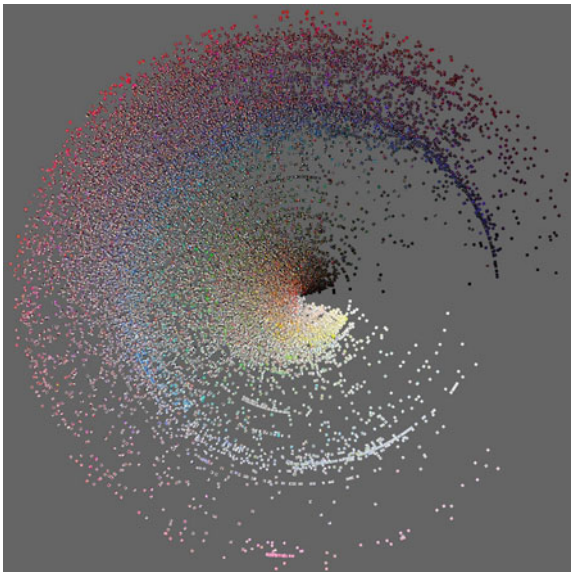

Fig. 2.4 From *Phototrails* project by Nadav Hochman, Lev Manovich, Jay Chow (2013) (reproduced with permission)

information referring to different geographic areas and different moments in time (including future trends) from different data sources related to a variety of themes and targeting different users profiles.

- In order to thoroughly design and assess this sensing platform, we collaborated with researchers, practitioners and users coming from different domains (academia, industry, governments, NGOs, and the public sphere). We describe in Chap. 4 how this collaborative process unfolded across our research trajectory.
- Critical perspectives toward the use of this kind of analysis are also taken into consideration, including technical limitations and issues related to the digital divide and privacy.

In more practical terms, we aim at designing a tool that can be easily accessible for urban planners, designers, and citizens with no specific technical skills: a sensing platform that allows final users to perform different operations on geo-referenced social media data in an easy and intuitive way. Users will be able to directly access the sources, filter information by all of its attributes, display content on a layered map, add additional elements (such as comments or labels), and export the results in different interactive and static formats through a web-based visual interface. Design wise, the main architecture of the platform will let different customized interfaces be plugged into address the specific needs of the diverse stakeholders.

In the Chap. 3, the state of the art of data-visualization strategies, depicting geo-referenced territorial data, is reviewed. Our aim is to identify some key information that can give important hints for the design and implementation of our sensing platform.

http://www.springer.com/978-3-319-02194-2

Visualizing the Data City
Social Media as a Source of Knowledge for Urban Planning and Management
Ciuccarelli, P.; Lupi, G.; Simeone, L.
2014, IX, 76 p. 24 illus., Softcover
ISBN: 978-3-319-02194-2