Urban Planning

Open Access Journal | ISSN: 2183-7635

Volume 3, Issue 1 (2018)

Crowdsourced Data and Social Media in Participatory Urban Planning

Editors

Bernd Resch, Peter Zeile and Ourania Kounadi

COGITATIO

Urban Planning, 2018, Volume 3, Issue 1
Crowdsourced Data and Social Media in Participatory Urban Planning

Published by Cogitatio Press
Rua Fialho de Almeida 14, 2º Esq.,
1070-129 Lisbon
Portugal

Academic Editors
Bernd Resch (University of Salzburg, Austria)
Peter Zeile (Karlsruhe Institute of Technology, Germany)
Ourania Kounadi (University of Salzburg, Austria)

Available online at: www.cogitatiopress.com/urbanplanning

Table of Contents

COGITATIO

Urban Planning (ISSN: 2183–7635)
2018, Volume 3, Issue 1, Pages 1–20
DOI: 10.17645/up.v3i1.1209

Article

Crowdsourced Quantification and Visualization of Urban Mobility Space Inequality

Michael Szell [1,2,3,4]

[1] Center for Network Science, Central European University, 1051 Budapest, Hungary; E-Mail: szellm@ceu.edu
[2] Center for Complex Network Research, Northeastern University, 02115 Boston, USA
[3] Complexity Science Hub Vienna, 1080 Vienna, Austria
[4] MTA KRTK Agglomeration and Social Networks Lendület Research Group, Centre for Regional Studies, Hungarian Academy of Sciences, 1094 Budapest, Hungary

Submitted: 2 October 2017 | Accepted: 18 December 2017 | Published: 29 March 2018

Abstract
Most cities are car-centric, allocating a privileged amount of urban space to cars at the expense of sustainable mobility like cycling. Simultaneously, privately owned vehicles are vastly underused, wasting valuable opportunities for accommodating more people in a livable urban environment by occupying spacious parking areas. Since a data-driven quantification and visualization of such urban mobility space inequality is lacking, here we explore how crowdsourced data can help to advance its understanding. In particular, we describe how the open-source online platform What the Street!? uses massive user-generated data from OpenStreetMap for the interactive exploration of city-wide mobility spaces. Using polygon packing and graph algorithms, the platform rearranges all parking and mobility spaces of cars, rails, and bicycles of a city to be directly comparable, making mobility space inequality accessible to a broad public. This crowdsourced method confirms a prevalent imbalance between modal share and space allocation in 23 cities worldwide, typically discriminating bicycles. Analyzing the guesses of the platform's visitors about mobility space distributions, we find that this discrimination is consistently underestimated in the public opinion. Finally, we discuss a visualized scenario in which extensive parking areas are regained through fleets of shared, autonomous vehicles. We outline how such accessible visualization platforms can facilitate urban planners and policy makers to reclaim road and parking space for pushing forward sustainable transport solutions.

Keywords
big data; bin packing; crowdsourcing; data visualization; mobility; OpenStreetMap; sustainable transport; transport justice; urban space inventory; volunteered geographical information

Issue
This article is part of the issue "Crowdsourced Data and Social Media for Participatory Urban Planning", edited by Bernd Resch (University of Salzburg, Austria), Peter Zeile (Karlsruhe Institute of Technology, Germany) and Ourania Kounadi (University of Salzburg, Austria).

1. Introduction

From a geometric perspective, urbanization is a process that packs large numbers of individuals into a limited amount of space. Over time, all urban space is zoned and allocated for human activities: dwelling, industry, business, recreation, and mobility. Due to this natural confinement and densification, urban public space has become highly contested, especially transport infrastructure, including roads, parking, sidewalks, tram tracks, and bicycle lanes (Gössling, 2016). Historically, between these mobility spaces, the car has been given highest priority, leading to car-centric cities, starting with its advent in the early 20th century (Norton, 2007). In the US, the initial medium in the struggle for transport space was language: new terms like "pleasure traffic", "joy rider", and "jay walker" shaped the public discourse, but by the 1930s, street equilibrium had been decided towards cars, paving the way for the "supremacy of automobiles" (Norton, 2007). Consequently, the automobile industry has systematically bought up and dismantled public street car systems (Urry, 2013) and demolished disadvantaged, usually black, neighborhoods to create space for cars with the help of public officials. One famous example includes the construction of the Eisenhower National System of Interstate and Defense Highways, which considerably improved travel connections between cities but destroyed and fragmented living space within them (Mohl, 2002). Thus, the evolution towards car-centric cities is a history paved with institutionalized racism and social injustice—a process that is repeated in non-Western parts of the world today (Martin, 2007).

1.1. The Spatial Inefficiency of Cars

Adding to the social injustice of car-centric policies, which are still in place today and continue to create divides between people who can and cannot afford a car, cultivating the automobile as a main mode of transportation has a number of sustainability issues. We first highlight those issues that relate to mobility space.

Due to individualistic, wasteful use of resources, privately owned cars are not used between 95% and 97.5% of the time (Bates & Leibling, 2012; Organisation for Economic Co-operation and Development [OECD], 2015; Weigele, 2014). For example, privately owned cars in Berlin are used on average 36 minutes per day. In other words, they are not used for 1404 minutes per day, or 97.5% of the time (Weigele, 2014). These low usage rates imply that at a typical point in time, out of the 1.2 million registered cars in Berlin, only 30,000 are actively used on the road, while the rest is parked. These parked cars take up massive urban space amounting to around 14 million square meters, the area of 64,000 medium-sized playgrounds (15m $\times$ 15m) or over 4 Central Parks. The same argument holds for London (Bates & Leibling, 2012) and other cities (Shoup, 2005). Further accessible estimations of these massive areas include: the total of all off-street parking spaces of the USA is approximately the size of Connecticut (Shoup, 2005); 7 million front plots in Britain have been replaced by parking lots, amounting to around 100 Hyde Parks (Bates & Leibling, 2012). Such inefficiency comes with exorbitant economic burden—Shoup (2005) estimates free parking in the US to correspond to a yearly public subsidy to car drivers of at least $127 billion.

Cars have an average occupancy of 1.5 individuals (Moriarty & Honnery, 2008). With a dimension of around $8m^2$, a car thus requires $5m^2$ per person, having a much higher mass-to-seat ratio than bicycles or well-utilized mass transportation. However, the striking space inefficiency of cars comes from heightened demands on infrastructure space, especially parking space (Chester, Fraser, Matute, Flower, & Pendyala, 2015; Shoup, 2005), and from secondary effects such as increasing transport volumes due to urban sprawl (Banister, 2005; Glaeser & Kahn, 2004; Gössling, 2016; Hutton, 2013). Due to these issues, car-centric cities have decreased benefits from scaling (Louf & Barthelemy, 2013).

Because automobile infrastructure takes up an excessive amount of urban space, we ask: Do we even know how exactly mobility space is distributed today? If not, can we measure how much more space is allocated to cars than to more sustainable forms of transportation? Is there an "arrogance of space" (Colville-Andersen, 2013), and how can we quantify and help to revert it?

1.2. Towards a Rigorous, Large-Scale Assessment of Mobility Space Inequality

Although the car-centricity of today's cities is a leitmotif in the urban transport planning literature (Banister, 2005), the distribution of urban mobility space and its "fairness" has so far not been quantified both rigorously and on a large scale. The quantification problem can be broken down into two steps: 1) Measure the distribution of mobility space between different modes of transportation; 2) Compare the share of allocated space with the modal share. If share of allocated space for one mode of transportation is substantially higher than its modal share, the distribution is unjust, because this mode gets more space than it "deserves" from actual usage.

A first attempt comes from Agentur für clevere Städte (2014), who have studied the distribution of space between cars and bicycles in 200 streets in Berlin. Agentur für clevere Städte found that only 3% of the streets have bicycle lanes, but 58% of the total transport space is allocated for cars. With a modal share of 15%, bicycles thus receive considerably less space, while cars, with a modal share of 33%, are vastly over-prioritized. At the same time, being faced with growing bicycle traffic, the study concludes with policy recommendations to extend bicycle infrastructure.

Another, more visual, approach comes from Colville-Andersen (2014) who has manually assessed the distribution of a few hand-picked intersections, and from visual inspection has indeed concluded an "arrogance of space", i.e., a privileged allocation of urban space to cars at the expense of sustainable forms of mobility such as cycling. Although this first attempt is visually impressive, its methodology does not follow a rigorous assessment of space and is not scalable. In his latest publicized case study, Colville-Andersen (2017) directly compares modal share with space allocation, and finds the imbalance between the two makes the issue particularly pressing: there is a 62% modal share for bicycles, but only 7% of mobility space is allocated to them. On top of this, the result is obtained for a particular street segment in Copenhagen which is known for its bicycle-friendly planning culture. If the arrogance of space is already so apparent in Copenhagen, then how bad is the situation in cities of infrastructurally underdeveloped countries like the US?

Gössling, Schröder, Späth and Freytag (2016) have recently quantified the inequality of urban mobility space scientifically. This study also focuses on a few handpicked intersections as proof of concept in the German city of Freiburg, but uses more rigorous methods involving satellite images and validation through high-resolution maps and on-site visits. Like Colville-Andersen (2017), Gössling et al. (2016) find that in their case studies far more space is given to streets and car parking than to bicycle lanes. Similarly, they conclude that pedestrians, on the other hand, receive ample space when compared to their modal share. It is important to note that having space is only a physically necessary, but not a sufficient prerequisite for a form of mobility to work successfully; making cities walkable or bikeable requires a number of conditions to be fulfilled (Speck, 2013).

In this article, we shift the approach from such individual, manual, and thus costly assessment of the distribution of mobility spaces (Colville-Andersen, 2014, 2017; Gössling et al., 2016) to data-driven, crowdsourced measurement. This approach has the benefit of being large-scale and inexpensive: it is able to capture *entire* cities due to the automatized computation of all mobility spaces that have been tagged by millions of users. On the other hand, this method comes with the disadvantage of less rigor than Gössling et al. (2016) due to reporting biases and data quality issues, leading to less accuracy on particular places. Nevertheless, the strength of the method lies in numbers: the statistical accuracy of assessing space inequality *increases* with scale, while the cost is relatively negligible and scales efficiently with the number of assessed square meters—whether we consider a neighborhood, a city, or a whole urban agglomeration. The focus of this article, however, is not to establish increased correctness in the assessment of mobility space distributions, but to explore advances in visualization, public engagement, and crowdsourced urban planning.

In this article we first present a novel method of collecting and visualizing city-wide mobility spaces for public exploration through an online platform in Sections 2 and 3. This presentation is followed by an analysis of data collected from visitor interactions on the platform in Section 4, providing evidence for a biased perception of mobility space inequality. In Section 5 we add a particular note on how much parking space could be regained if all current cars turned into a shared, public fleet of autonomous vehicles—ignoring the feasibility and unintended consequences of such a scenario. We discuss possible data and design limitations of crowdsourced data and online platforms in Section 6, improvements and synergies between the different assessment methods and their relevance for urban planning in Section 7. Further sustainability issues of cars and their possible solutions are discussed in Section 8.

2. Visualization of All Parking and Lane Spaces

To visualize the space requirements of different forms of mobility, with moovel lab we recently developed and launched an interactive online platform, *The Mobility Space Report: What the Street!?*[1]. The idea of the platform is to collect all mobility and parking spaces of a whole city for each type of mobility, and to align these spaces in a visually comparable way. At the same time, the platform makes the "arrogance of space" accessible to a broad public, packing all urban mobility spaces into giant bar charts, reducing the complexity of comparing thousands of shapes to the single dimension of comparing heights. Data for parking and lane spaces originate from OpenStreetMap (OSM), the crowdsourced open-content alternative to commercial online maps. Due to data availability reasons (see Section 6) only three types of mobility were considered: cars, rails, bicycles.

2.1. Data Collection, Processing, and Selection

For each form of mobility the platform deals with two sets of spaces: 1) parking spaces; and 2) spaces that are used for movement. In the case of automobiles and bicycles, the parking spaces are encoded by polygons. For all three forms of mobility, the spaces that are used for movement are given by polygonal chains (curves specified by sequences of points) and an optional width. Rail parking consists of service tracks, also encoded by polygonal chains.

All methods of data collection, data wrangling, and the technical implementation of the visualization, including the complete code for back and frontend, are documented and open-sourced.[2] The whole process is thus completely reproducible, and summarized in the following paragraph. Further technical details can be found in the repository READMEs.

Data collection was a straightforward download, either directly from OSM, or from a content aggregator like Geofabrik.[3] The geo-data was cropped with the city limits using the OSM-specific tool *osmconvert*. The cropped data was then loaded into a MongoDB. The raw OSM data consist of a multitude of elements, including nodes, ways, and relations such as traffic lights, forests, or restaurant locations. To filter this data for the relevant car and bike parking structures, the corresponding polygons were identified using the *amenity=parking* and *amenity=bicycle_parking* tags, respectively. For car lanes, i.e. roads, all street names were first identified, to only select roads that are searchable by name. Using these names, road spaces were then selected via *highway=service OR highway=residential OR highway=primary OR highway=secondary OR highway=tertiary OR highway=unclassified*. Bicycle lanes were selected to include only physically separated lanes,

[1] https://whatthestreet.moovellab.com
[2] https://github.com/moovel/lab-what-the-street
[3] www.geofabrik.de

i.e. lanes that follow best practice implementation (see Section 6.5) using the tags *highway=cycleway OR bicycle=designated OR cycleway=track*. Rail space used for movement was selected with the tags *railway=mode AND (service=crossover OR !service)* where mode is one of *tram*, *light_rail*, *rail*, *subway*, *narrow_gauge*, *funicular*, *monorail*. Parking rail spaces were selected using *railway=mode AND service!=crossover AND service*. A parser was developed to extract all relevant information on number of lanes and widths accounting for inconsistent or ill formatted data, to identify areas correctly. For details on data selection and biases see Section 6.

2.2. Parking Space Packing

Parking spaces come in a variety of scales and shapes, from rectangular lots for single cars to meandering structures that accommodate many thousand vehicles. These shapes are encoded as polygons in OSM. A straightforward way to pack these spaces densely into a rectangular bin is the application of a polygon packing algorithm, solving heuristically the irregular bin packing problem (López-Camacho, Ochoa, Terashima-Marín, & Burke, 2013). The platform uses the open-source package *SVGNest*,[4] which was initially developed for minimizing waste when cutting shapes out of a flat material. Figure 1 (top) shows *SVGNest* applied to all car parking spaces of the city of Johannesburg (rotated by 90 degrees). Due to restrictions on computational complexity—bin packing is NP-hard (López-Camacho et al., 2013)—*SVGNest* ran iteratively on randomly partitioned, then re-stitched, subsets of the polygons, fully exhausting them. These technical steps are documented in the corresponding linked source code. The purpose of this auxiliary heuristic is to overcome computational limitations of *SVGNest* while balancing aesthetics.

Visually, this process leads to the aesthetically pleasing and almost seamless packing of all parking spaces as demonstrated in Figure 1 (top). This rectangle of packed parking spaces can then be directly compared to another rectangle of mobility spaces if it contains objects with the same density.

2.3. Lane Cutting and Coiling

For each city and mobility type the platform uses all lane spaces. To make lane space comparable to the corresponding parking space, lanes were coiled and stitched together in a sausage-like tube, shown in Figure 1 (bottom). The displayed thickness is not to scale but was chosen for aesthetic consistency. However, the underlying density of the tube was calculated using the weighted average of all lane widths in the city, which gives it the same density as the corresponding parking space bar, making these two bars directly comparable through their heights.

In the case of rails and bikes, a lane is encoded with the "way" data structure of OSM. Ways are polygonal chains, which makes rail and bicycle lanes homeomorphic to their coiled counterparts. In other words, they can be bent smoothly without the need for cuts or stitches. Car lanes, on the other hand, were defined by street name to be searchable. Streets are typically not simple polygonal chains, but multiple polygonal chains stitched together, representing trees or even more complex graphs that contain cycles. For example, Figure 2 shows a street with branches. Furthermore, usually there are multiple streets in a city with the same name. Therefore, to handle the coiling of streets, the platform was equipped with an algorithm for the non-continuous transformation of each street graph into a chain of polygonal chains that can be bent smoothly. The algorithm iterates through all connected components; each component is traversed via Depth First Search. This step is repeated until no piece is left, taking care to store a minimal data structure during traversal which allows correct reconstruction of all pieces.

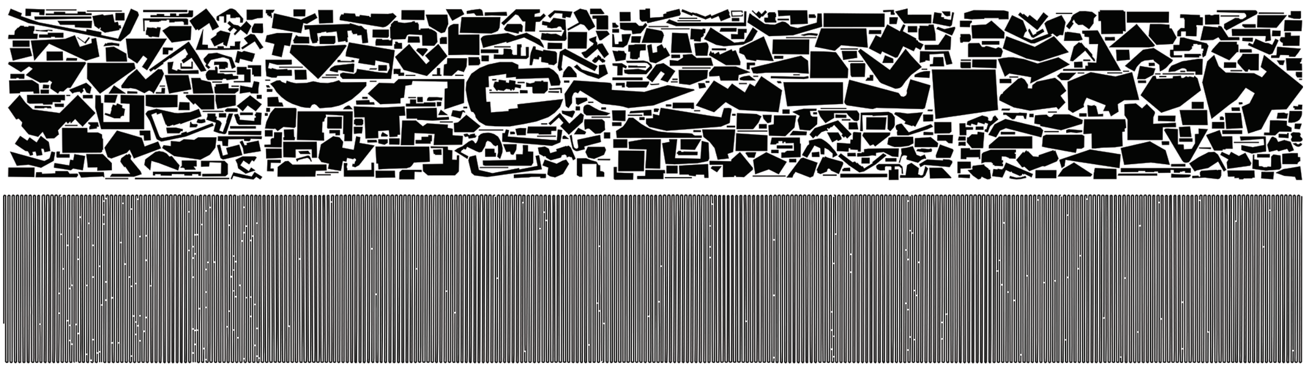


Figure 1. Top: All OSM car parking spaces of Johannesburg packed into a rectangle using the polygon packing library SVGNest. Bottom: The matching area of a segment of coiled up OSM streets of Johannesburg. The figure is rotated by 90 degrees.

[4] http://svgnest.com

Figure 2. How a street uncoils: a. coiled view on map; b. zoomed out; c. unrolled; d. split up and translation of misplaced parts that previously exceeded the viewport in the bottom right corner—this is the original shape of the street.

2.4. Individual Exploration and Animation

As an additional way to engage visitors, all packed and coiled objects are interactive and individually explorable. Figure 3 illustrates the main interface, with car parking spaces packed on the left bar and streets coiled on the right. Information is given on the top center on the total amount of so far scrolled area, here 419,489 m^2, in a human readable way, of 2.2 Boston Common landmarks. An individual landmark was selected for each city, e.g. Central Park in New York City or Hyde Park in London. Scrolling along the bars highlights individual elements on each bar, providing additional information: neighborhood (East Boston), street name (Appleton Street), and area. When clicking on a parking space, the polygon appears on a map and rotates from the packed minimum-width orientation into its correct map orientation, displaying information on how many cars or bikes can be accommodated approximately. When clicking on a lane, the coiled lane appears on the map and uncoils into its original position on the map (Figure 2), displaying information on length and area of the lane. If the street is not made up of a single polygonal chain, only one component is now in place, while all other components are attached to one end being misplaced (Figure 2c). In this case, a second step in the animation translates the misplaced pieces, ensuring that all pieces of the street return to their original place on the map (Figure 2d).

3. Visual Assessment of Space Imbalance Using the Mobility Triangle

Parking spaces are packed and lane spaces are rolled into bars that can be compared to each other by scrolling

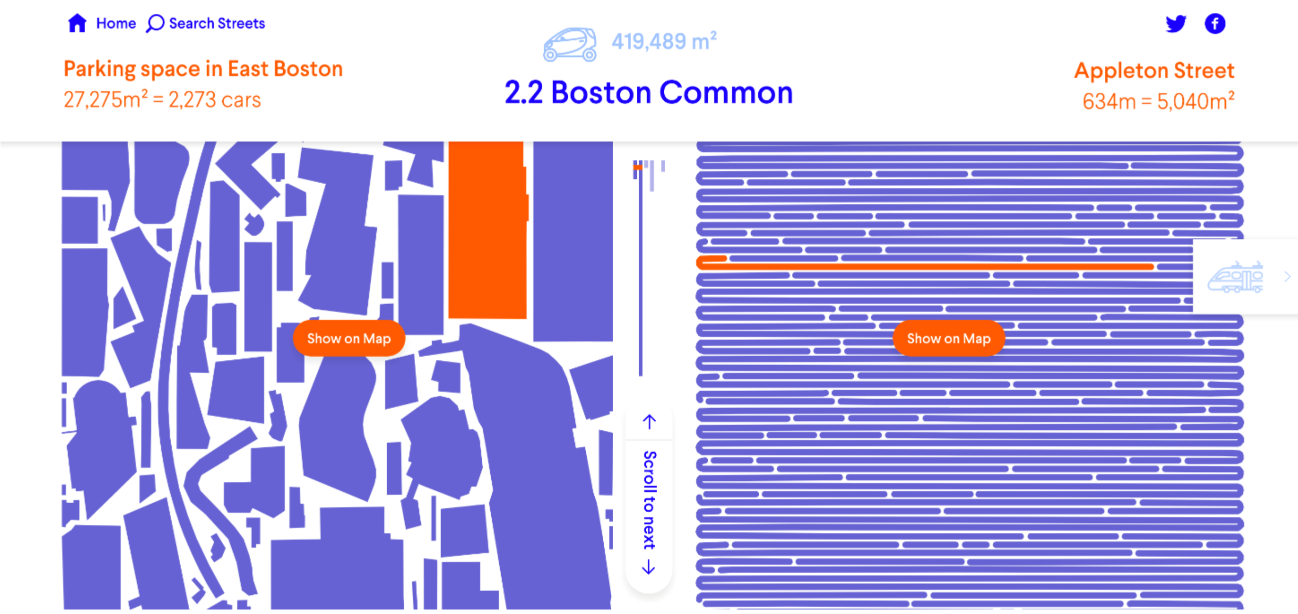


Figure 3. The main interface allows scrolling through all mobility spaces, here shown for cars in the city of Boston: on the left a segment of the bar of packed parking spaces, on the right the coiled streets with all individual elements explorable in detail (one on each side highlighted in orange).

through them. The time needed to scroll through the vast car spaces, in contrast to the rather short bike spaces, gives an intuitive, dynamic perception of their inequality. However, there is a more compressed form of comparing these spaces to each other that does not require scrolling or comparing bar heights.

Each city has its own distribution of total car to rail to bicycle space, given by a triple (X, Y, Z) where $X+Y+Z=1$. These variables identify the Cartesian coordinate (x,y) with $y = Y*\sin(\pi/3)$ and $x = X + y*\cot(\pi/3)$, encoding the distribution as a single point in a ternary plot (Figure 4). A dot in the top corner of this triangle means all mobility spaces are car spaces, a dot in the bottom left corner means they are all rail spaces, and a dot in the bottom right corner means they are all bicycle spaces. A dot in the center of the triangle means that an equal share of space is allocated to each of these three forms of transportation.

The same coordinate transformation can be performed for the city's modal share, yielding a second data point. Connecting both dots in the ternary plot yields the "mobility triangle". It shows with just two data points the discrepancy between allocated space distribution and how people actually move. If first and second dots coincided, it would mean that city space is allocated in a fair way towards all forms of mobility: cars, rails, and bicycles receive the same fraction of space that they "deserve" from their usage. To give some real-world examples, Figure 4 shows the mobility triangle for the cities of Beijing, Berlin, Budapest, Copenhagen, Los Angeles, Rome, and Tokyo. Apparently, the dots representing allocation of space are clustered toward the top car corner of the triangle, while the modal share dots are spread towards the bottom and to the right, visualizing how many more people use other forms of mobility than is allocated for cars. We chose these specific 7 cities out of the 23 available ones[5] for Figure 4 because of the visual separation of their data points, and because they cover the most extreme cases (modal share dominated by rails, bicycles, or cars).

For comparison, we present the same data in the form of a table (Table 1). The stark contrast in accessibility highlights why visual communication is far superior for numerical data, an insight established since at least the 19th century:

> The graphical method has considerable superiority for the exposition of statistical facts over the tabular. A heavy bank of figures is grievously wearisome to the eye, and the popular mind is incapable of drawing any useful lessons from it as of extracting sunbeams from cucumbers. (Farquhar & Farquhar in Wainer, 2005, p. 9)

4. Empirical Evaluation of Space Inequality Perception

Before engaging platform visitors into any of the above interactive visualizations, the platform invites them on its

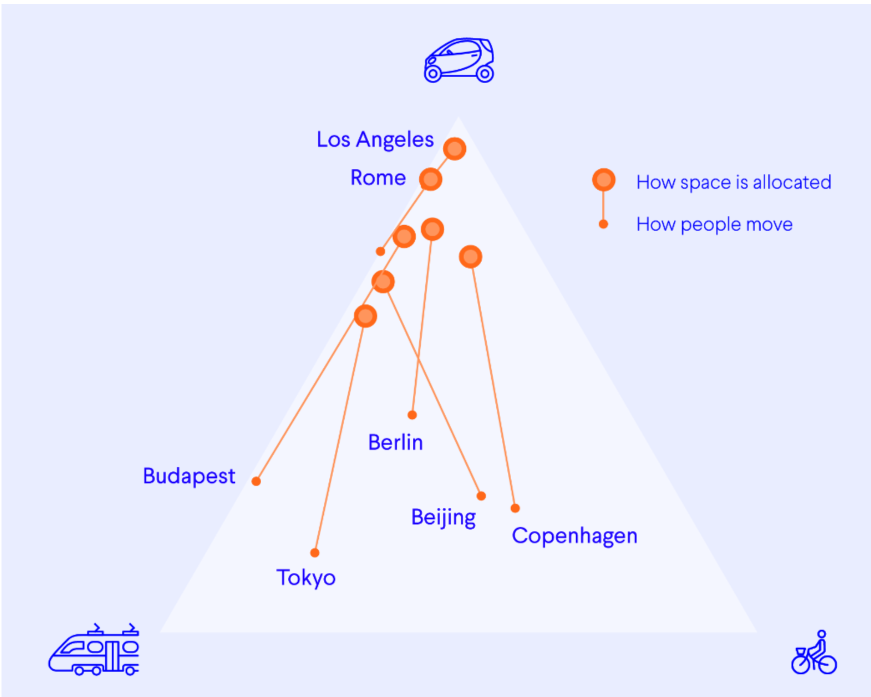


Figure 4. The mobility triangle expresses visually the discrepancy between mobility space distribution ("How space is allocated") and modal share ("How people move") as a connected pair of dots in a ternary plot, shown for seven selected cities.

[5] The mobility triangle is shown online for all 23 cities on the platform page, though with higher visual clutter. To avoid such visual clutter we recommend to limit the number of cities displayed in the same mobility triangle plot to not more than 8.

Table 1. Mobility space distribution versus modal share in seven selected cities. Modal shares do not sum up to 100 because of walking which accounts for the remainder. Some space distribution rows do not sum up to 100 due to rounding.

	Space Distribution (%)			Modal Share (%)		
City	Automobile	Rail	Bicycle	Automobile	Rail	Bicycle
Beijing	68	29	3	21	26	32
Berlin	78	15	7	30	26	15
Budapest	77	21	3	20	47	1
Copenhagen	73	12	16	23	27	45
Los Angeles	92	4	4	78	11	1
Rome	88	11	1	68	24	0
Tokyo	61	35	4	12	51	14

landing page to make a guess, asking for their estimate on how much space is allocated to the different forms of mobility. This invitation is formulated via two questions: "Who owns *cityname*?" (Figure 5), and "City space is limited! What do you think, how much space is allocated to the different ways of moving through the city?", where *cityname* stands for the name of one of the currently 23 covered cities,[6] and is automatically pre-selected to be closest to the visitor's location, estimated by IP address. The guess is made with a set of 3 sliders linking cars, rails, and bicycles, adding up the percentages to 100%. A guess is recorded only if the individual moves any of the linked sliders and makes her choice later than 8 seconds after having arrived on the page.

The platform collected 2,436 guesses from people across the 23 cities between the 4th and 25th of July 2017; the results are reported in Figure 6. Box plots visualize the guesses versus dot markers that show space allocation as determined by the OSM data. The data show that people consistently overestimate the space given to bicycles, and underestimate the space given to cars. In 22 of the 23 cities, the median of guesses for bicycle space lies above the fraction as determined from the OSM data. The exception is Helsinki; however, here we suspect an issue with OSM data or the way the platform is parsing it, since an allocated bicycle area of 26%, much more than the 16% in Copenhagen and 19% in Amsterdam, seems unrealistic. We measure a similar, but inverted picture for the perception of car space. In 22 of the 23 cities, the median of guesses for car space lies below the fraction as determined from the OSM data. The only exception is Vienna, where the median of guesses is 67.5%, slightly above the OSM fraction of 63%.

There are possible biases which call for a careful interpretation of the significance of these results. First, it is not clear how the initial configuration of the slid-

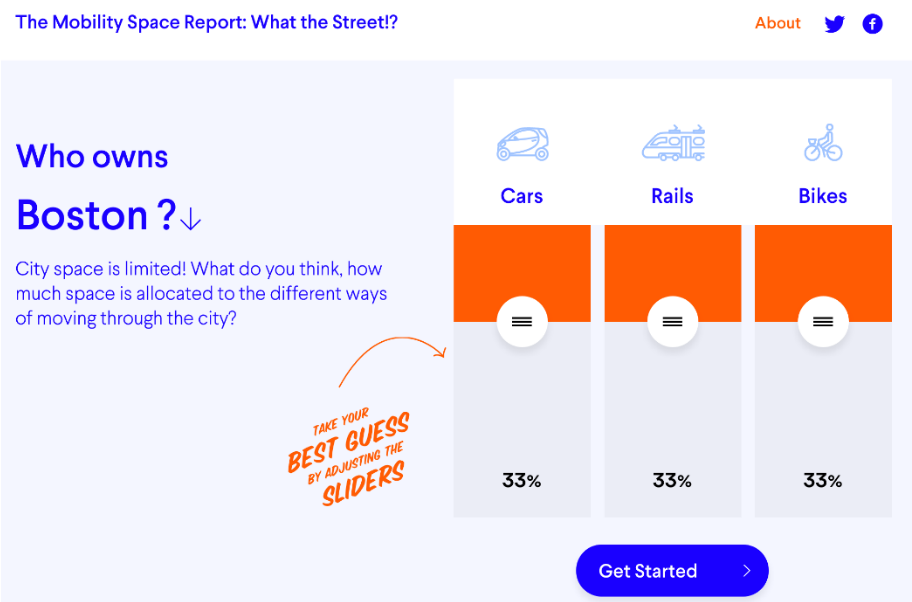


Figure 5. Using a linked set of sliders, empirical guesses on the distribution of mobility space are collected on the landing screen from site visitors.

[6] Amsterdam, Barcelona, Beijing, Berlin, Boston, Budapest, Chicago, Copenhagen, Helsinki, Hong Kong, Jakarta, Johannesburg, London, Los Angeles, Moscow, New York City, Portland, Rome, San Francisco, Singapore, Stuttgart, Tokyo, Vienna.

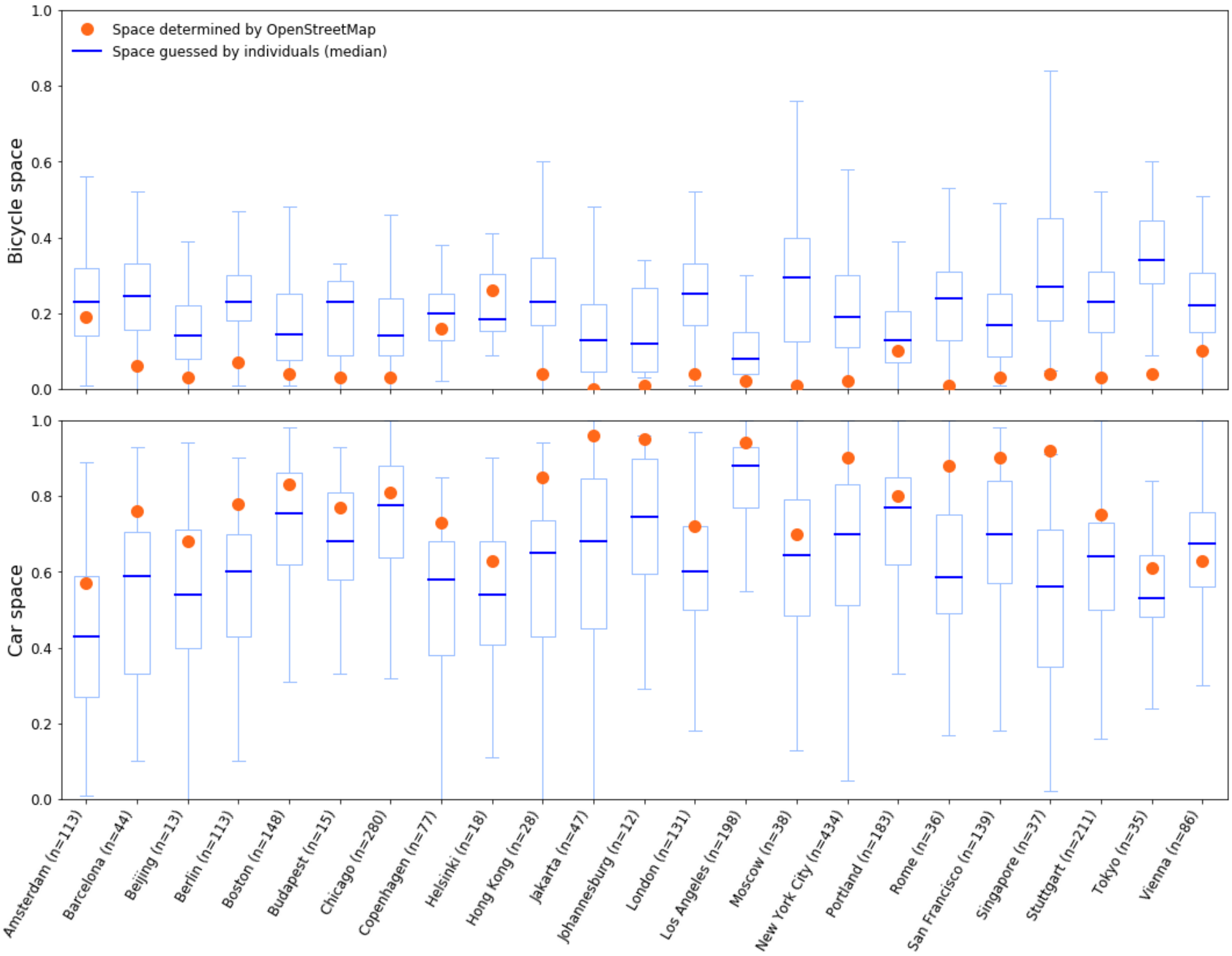


Figure 6. Guesses collected show that people consistently overestimate the space allocated for bicycles (top) and underestimate the space allocated for cars (bottom).

ers at 33%–33%–33% or the order "Cars-Rails-Bikes", Figure 5, biases the guesses. It is possible that this setup shifts choices towards a more homogeneous distribution due to anchoring bias. There is survey-methodological research on slider bias versus radio button bias (Couper, Tourangeau, Conrad, & Singer, 2006), however with unclear relevance since the platform uses linked sliders. We are not aware of survey bias research on the rarely used linked sliders. Second, it is not clear how the loaded question "Who owns *cityname*?" biases guesses. Third, most of the individuals who arrived at the platform were referred to from media sites that covered the spatial inequality aspect, mostly in English or in German. The individuals in this biased sample are therefore likely already sensitive to transport injustice and should tend towards a more informed guess than average internet users. Fourth, on top of this sample bias, it can also be inferred that most of the site visitors are WEIRD (Western, educated, and from industrial, rich, democratic countries) because of the number of guesses being heavily skewed towards Western locations: $n > 100$ guesses are made for each US city and for several European cities, but there are only $n \ll 100$ guesses for the other cities. (All sample sizes are shown in brackets in Figure 6.) We do not know how many of the site visitors are locals, i.e. living in the city about which they make their guess, or if they make guesses only in their IP-assigned or in multiple cities, but we assume that whenever they make one or more guesses then at least one of these is in their assigned city. The differences in guesses between different cities suggests that a substantial fraction of the individuals is not guessing independently of the city, but consists of either locals or is sensitive towards differences in space allocation. For example, Los Angeles, a well-known sprawl city, has a median guessed car space of 88% (±14% SD), while this guess is less than half, 43% (±21% SD), for Amsterdam which is known for its bicycle culture. The correlation of guessed medians versus OSM areas over all cities is $\rho = 0.63$ (p=0.0013) for cars. Interestingly there is no significant correlation for bicycle spaces, $\rho = 0.04$ (p=0.84), suggesting that most individuals judge a city by its car space (which might be influenced by the slider order "Cars-Rails-Bikes"). The strong correlation for cars shows that people consistently underestimate car space *even though* they adjust their guesses towards the real situation.

To sum up, there are possible biases that could have skewed the results in both directions, and the differences between guesses and space allocations are not statistically significant for any single city, in terms of mean ± 2 SD, due to the high variation of guesses. However, the consistent underestimations of car spaces across the board means that the overall result is significant: the probability for the null hypothesis of underestimating 22/23 cases by random chance is $23*0.5^{23} < 0.001\%$. Therefore, we interpret our empirical observations as adequate evidence that the issue of unequal space allocation between modes of mobility is underestimated by most people. This result makes the issue of urban transport injustice even more pressing, as car space is underestimated by the general public. (Further research is however needed to confirm this misperception and to understand how it effects urban planning processes and transport policy decisions either directly by biased decision makers or indirectly through a biased population that lacks motivation to assert political pressure towards sustainable transport solutions.)

5. Exploring a Scenario on Parking Space Reduction

The last section of the platform explores the scenario in which all privately owned cars are replaced by shared, self-driving vehicles. Due to much more efficient use of the vehicles, the fleet is assumed to be reduced by 90% (OECD, 2015) with parking spaces reduced by 93% (Friedrich & Hartl, 2016). For each city, the platform gives key statistics and a graphical impression on total saved landmarks, from Vondelparks in Amsterdam to Schlosspark Schönbrunns in Vienna (Figure 7). Without information on on-street parking, the numbers shown only account for a city's registered vehicles and are thus vastly underestimating the total reduced parking area—for some US cities at least 3.3 to 7-fold (Chester et al., 2015; Shoup, 2005). Further underestimations come from differences in counting parking spaces. For example, Chester et al. (2015) find that area devoted to parking is actually 1.4 times larger than the total area of roads in Los Angeles county. This discrepancy comes from methodological differences: they count individual parking spaces, including multi-story parking garages, while the platform only considers surface area taken up by parking. The methodological underestimation of the platform, apart from data limitations, underlines the huge amount of space devoted to road infrastructure and difficulties in assessing the actual scale of it.

Although this particular scenario is based on simulations in optimal settings, without discussing feasibility nor possible unintended consequences, the visualization provides a proof of concept how crowdsourced online platforms could be useful for accessible urban scenario planning (Peterson, Cumming, & Carpenter, 2003). Note that a more recent study using data-driven simulations in

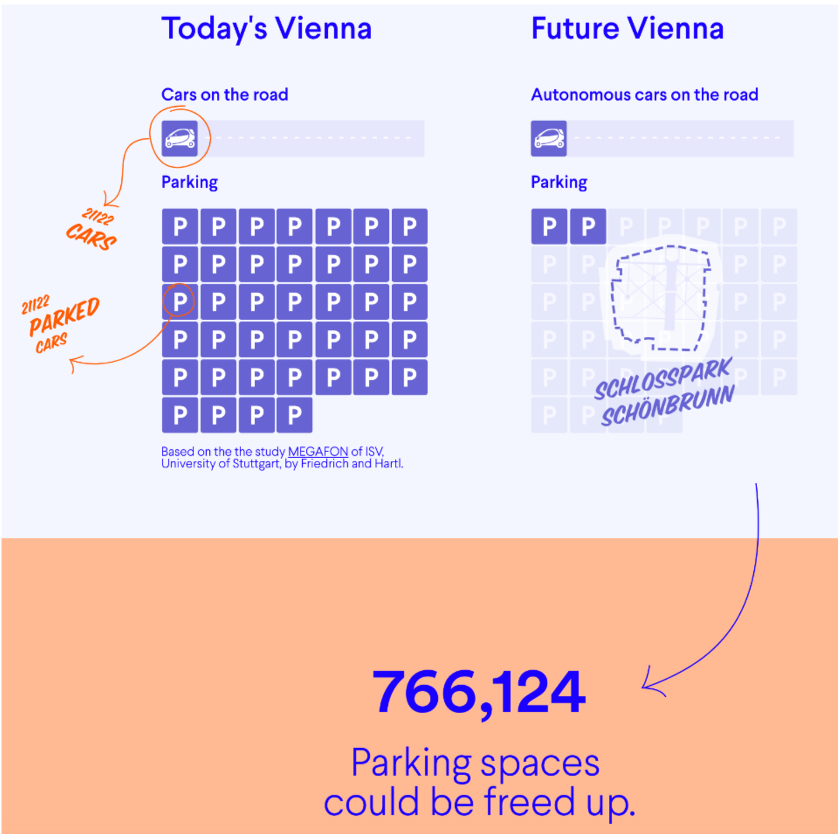


Figure 7. The spatial consequences of a scenario in which all cars in a city are replaced by an autonomous, shared fleet of vehicles.

Singapore reports parking space reductions through self-driving vehicles in the order of only 50%, at the expense of increasing total traveled kilometers by less than 2% (Kondor, Zhang, Tachet, Santi, & Ratti, 2017). As quantitative research on parking spaces is extremely sparse despite its fundamental role in city planning (Chester et al., 2015; Shoup, 2005) further research on the topic is urgently needed.

6. Data and Design Limitations

The biggest limitation of crowdsourced platforms is their data source, in this case OSM, and all the biases and shortcomings that these sources entail.

6.1. Data Quality

OSM is one of the largest examples of volunteered geographical information (VGI) today, having over 2.8 million registered users and over 3.4 billion contributed geographic data points (Zhang & Malczewski, 2017). Although there are research gaps in defining appropriate quality indicators, the VGI data quality literature focuses on the following main measures: completeness, consistency, and positional accuracy (Senaratne, Mobasheri, Ali, Capineri, & Haklay, 2017).

The main completeness issue of VGI systems such as OSM is the heterogeneity of users and records due to the digital divide. In particular, Haklay (2010) demonstrated the lack of coverage in non-Western parts of the world, and within each region in rural and poorer areas:

> Places that are perceived as 'nice places', where members of the middle classes have the necessary educational attainment, disposable income for equipment, and availability of leisure time, will be covered. Places where population is scarce or deprived are, potentially, further marginalised by VGI exactly because of the cacophony created by places which are covered. (Haklay, 2010, p. 700)

This heterogeneity is also a major source of discrepancy between OSM and governmental bodies like the Ordnance Survey that operate on the principle of universal service. Haklay (2010) found, for example, a high variation in the coverage rate of English roads—between 46% for the poorest, and 76% for the wealthiest neighborhoods. Variations in completeness also stem from a heterogeneity of users. OSM contributions are made by both power users and occasional users, who produce no bias in terms of content, but a bias in meticulousness that varies from culture to culture (Quattrone, Capra, & De Meo, 2015).

The issue of positional accuracy comes down to the same argument. Haklay, Basiouka, Antoniou, and Ather (2010) argue that Linus' law applies to accuracy in the case of OSM: "Given enough eyeballs, all bugs are shallow" (Raymond in Haklay et al., 2010, p. 70). Haklay et al. (2010) demonstrate in a quantitative example an inverse, nonlinear relationship between the number of contributors and average positional accuracy of the English road network: 15 or more contributors can improve accuracy to below 6m from a single-contributor accuracy of 11m. Comparison between the OSM and the ordnance data shows a quite accurate 80% overlap of motorway objects between the two datasets. However, this result refers only to covered areas, and again the problem of completeness dominates. Because of this and further issues, Zhang and Malczewski (2017) conclude that Linus' law is not applicable in VGI in general.

Given these issues, any rigorous assessment of mobility spaces, as collected from OSM by a visualization platform has to proceed with utmost caution. Absolute areas are most certainly not comparable between different cities. For example, Jakarta has an area of 661.5 km^2, and in OSM a total car parking area of 1.37 km^2. On the other hand, Singapore, covering a similar landmass of 719.1 km^2, has an OSM total of 4.23 km^2 car parking space. It is not clear whether this 3-fold discrepancy is rooted in a lack of OSM users in Jakarta, in the way OSM users record data, or in actual differences—OSM data of Singapore contains a few extended parking spaces and a large number of small-sized ones, while Jakarta does not contain many small spaces.

If absolute areas are not comparable globally, how can crowdsourced geographical information be useful to assess mobility space inequality? We argue that, despite the absolute differences between different cities, the relative differences within a city should be more stable. Apart from possible cultural biases (Quattrone et al., 2015), it seems reasonable to assume no significant difference in the biases in which the areas of car space, bicycle space, or rail space are recorded by users within the same city. Given this assumption, the central limit theorem implies that with enough spaces sampled from these three sources, the collected sample distribution converges towards the distribution of all spaces, meaning that the car to bicycle to rail areas sampled by OSM will closely resemble the true distribution. However, further research is required to assess the influence of user heterogeneity and how VGI co-evolves with modal share and urban space.

6.2. Missing On-Street Parking by Design

OSM does not record on-street parking in a polygonal format, but rather includes on-street parking information using *parking:lane* attributes within highway features. Because of this limitation, the area and shape of on-street parking is not available by design. In fact, cities themselves often do not have an up-to-date inventory of their parking spaces, and researchers are dependent on manual sampling methods. Following such an approach, Weinberger, Seaman, Johnson and Kaehny (2008) have shown that on-street parking can be a considerable fraction of all parking in a city. Therefore, by only considering off-street parking through the *amenity=parking* tags, an OSM based tool is likely vastly under-estimating the total car parking area within a city. In case that an OSM street contains a width attribute, it is not clear whether the user who added the tag intended to include on-street parking space; however, we assume this not to be the case because OSM specifies on-street parking information to be dealt via the *parking:lane* attribute. In case of no width attribute, the platform uses lane information and approximate fallback values for default street widths that do not contain on-street parking.

6.3. Missing Pedestrian Spaces

Although OSM has a *sidewalk* attribute it seems to be rarely used. Because of this lack of data, the area and shape of pedestrian mobility spaces is not available, apart from explicit off-street footpaths. Further, footpaths have no parking equivalent like the other forms of transportation, so the design of the platform would break. However, adding pedestrian spaces through other data sources (see Section 7), would be desirable to understand how motor-dependent or how walkable a city is.

6.4. Overlaps in Mobility Space

Different modes of transport often use overlapping spaces. For example, buses usually use the same road space as cars, bicycles use road space when there are no protected bicycle lanes, or tram rails might overlap with car space. Therefore, a unique attribution of each square meter of mobility space to one form of transportation is not possible. We followed a few simplifying steps that might distort the space inequality assessment. First, bus space was entirely ignored due to the substantial overlap with car space, and because it is not straightforward to quantify these spaces from the OSM data. Instead of a distinction into cars, public transport, and bicycles, space was split into cars, rails, and bicycles. Second, bicycle space was not considered whenever it neglected established practices of physical separation from car traffic as encoded by the *cycleway=lane* tag. Only physically separated bicycle space was considered, using the tags *highway=cycleway*, *bicycle=designated*, and *cycleway=track* (see Section 6.5). Third, possible overlaps of rail with car spaces were ignored. It could be argued that rail infrastructure competes less for mobility space than overlapping spaces do such as car and bicycle spaces. A possible alternative version of the mobility triangle could therefore replace rail space with pedestrian space.

6.5. Protected versus Unprotected Bicycle Space

The platform only considers protected bicycle lanes, i.e. lanes that are physically separated from vehicular traffic. Although it is outside the scope of this article to analyze bicycle infrastructure and cyclist injury rates, this choice is important and deserves an explanation. We first refer to the literature on bicycle infrastructure safety to argue why this is a reasonable choice, and then show that even if the platform were to consider both protected and unprotected bicycle lanes, it would not make a relevant difference—both the mobility triangle and the results on infrastructure perception would stay qualitatively the same.

Researchers who study bicycle infrastructure and injuries report growing evidence on two necessary ingredients for making bicycle infrastructure safe: 1) physical separation from vehicular traffic, and 2) improved intersection design. For example, Teschke et al. (2012) study injury risk in Canadian cities using a case-crossover design for statistical control via within-route randomization, finding a 9 times lower injury prevalence on physically separated cycle tracks than on reference routes. Shared bicycle infrastructure shows no significant risk reduction. Reynolds, Harris, Teschke, Cripton and Winters (2009) review 23 papers on the topic and report that roundabouts in particular can pose an increased risk to cyclists if cycle tracks are not separated. Pucher and Buehler (2016) review 8 recent studies and conclude that decades of evidence in Europe and the US demonstrate clear evidence that it is "crucial to provide physical separation from fast-moving, high-volume motor vehicle traffic and better intersection design" (p. 2090). These insights are not only academic, but increasingly are incorporated into policy guidelines such as the National Association of City Transportation Officials (2017) guidelines, which now recommend protected bicycle lanes as best practice except for the special circumstance of streets with low-speed, low-volume motor vehicle traffic. It therefore makes sense to not consider bicycle space that is accessible by cars, as it cannot be considered safe for cyclists in typical traffic conditions. Despite all the empirical evidence found so far in favor of protected bicycle lanes, cycle tracks are perceived less safe by the public than observed, while multiuse paths are perceived safer than observed (Winters et al., 2012). Further, since the OSM data do not allow to distinguish between properly and improperly designed intersections and lanes, the OSM protected bicycle lanes can only be considered a proxy for best practice implementations.

Because bicycle infrastructure safety literature is still sparse, let us ask: Are these results robust to the issue of protected versus unprotected bicycle lanes? To understand this question, we counted all unprotected bicycle ways via *cycleway=lane OR cycleway=opposite_lane OR cycleway=share_busway* from OSM and calculated how much their count increases from the previously considered count of protected bicycle lanes. We find that 12 of the 23 cities add only less than 10% bicycle ways if also unprotected ones are considered (Hong Kong 0%, Helsinki 1%, Singapore 1%, Moscow 2%, Tokyo 2%, Rome 2%, Amsterdam 3%, Stuttgart 4%, Jakarta 5%, Beijing 6%, Copenhagen 7%, Barcelona 8%). Seven cities add between 10% and 50% (Johannesburg 15%, Portland 29%, Berlin 33%, London 35%, Chicago 36%, Budapest 39%, Vienna 42%). The final four cities are all from the US: Boston 73%, New York 160%, Los Angeles 182%, San Francisco 222%. Since these cities have an already negligible absolute area of bicycle infrastructure, doubling or tripling it does not make any qualitative difference in the mobility triangle (Figure 4), nor in the result on misperception (Figure 6). In light of the previous discussion of how most unprotected bicycle lanes could be considered unsafe, it is not surprising that the Western country with the strongest car culture has by far the highest ratio of unprotected bicycle space. This preliminary observation motivates further research on whether, how, or why societies that provide the least space for bicycles are also the societies that provide the most inadequate space for bicycles.

6.6. Overlaps in Modal Share

Modal share is the percentage of travelers commuting by walking, cycling, public transport, or private motor vehicle. There is an issue with this distinction in the context of the platform. As above, public transport incorporates buses, which share the same space as cars and cannot be uniquely separated. Further, public transport is a mix of bus and rail transport, which cannot be directly compared to the mobility space of rails because of the missing bus spaces. Therefore, a one-to-one mapping between mobility space and modal share is not possible, and the "How people move" dot in the mobility triangle (Figure 4) should be closer to the top corner if bus space is taken as car space. A general issue with modal share is its measurement via surveys by local governments without standardized methods, implying that modal shares of different cities have to be compared with caution. Here, the platform developers gathered modal shares from various official sources.[7] Regardless of the data source, the contribution of the platform lies in a new visualization tool, in particular the mobility triangle. The accuracy of this visualization can only be as good as the accuracy of the underlying data.

6.7. Web Design Limitations and Possibilities for Improvements

The platform was developed by a team of 15 people in total, coordinated by moovel Group GmbH, consisting of user experience designers, graphic designers, web designers, Javascript experts, mobility experts, and data scientists. This team had to balance web design, user experience, available web technology, and data limitations. For example, questions like "How many parking spaces can we fit into a browser window?" or "How do we load thousands of polygons into the browser and still make it scrollable smoothly?" had to be solved together. Because of many such constraints, the complexity of interactive parts, and a finite budget, there are still several possible avenues to improve the platform. For example, it could be optimized for mobile view, a scale or landmark could already be shown during the 2-column view, there could be a tutorial, the types of roads or rail tracks could be distinguished, cars and busses could be merged as a "motor" category, or many usability improvements could be implemented such as accent-insensitive search ("Nador utca" should return the same street as "Nádor utca"). Fortunately, because the platform is open-sourced, anybody is free to open an issue or to implement such improvements.

7. Potential Use in Urban Planning and Outlook

In the current form, the platform collects, quantifies, makes explorable, and summarizes with visual elements and key statistics the crowdsourced, massive data sets composed of millions of single mobility spaces from OSM spanning city limits. Primarily, it is useful as an educational tool for the public, to engage with the issues of sustainable transport and transport justice, and to make tangible how much parking space is wasted that could be utilized more meaningfully, through non-technical advances (policy changes) or technical advances (e.g. shared, autonomous cars). Such public engagement has the potential to improve public opinion and the measured misperception of sustainable transport and could indirectly back urban policy makers to reclaim road and parking space for pushing forward sustainable transport solutions (Chester et al., 2015). A public ranking of different cities could also come with positive competitive effects.

As a direct aid for urban planners, the platform gives a quantified overview of mobility space distributions spanning an entire city. Although inventoried data of such spaces may already be available to city administrations, the automatized merging and public processing provides a boost in accessibility. Because it is open-sourced with an MIT license, it can be applied to any city, and extended or adjusted to fit particular use cases as a visual quantification software. For example, longitudinal data could be collected to compare developments in time, to check if new policies are required or have an impact, or to identify underdeveloped neighborhoods in need of focused investments.

7.1. Longitudinal Extension to Study Induced/Disappearing Traffic

The discrepancy between modal share and available space should not be underestimated, but cast in a systems dynamics perspective. The ternary nature of the mobility triangle is a hint at an evolutionary game theory setting where three strategies are competing against each other: extend infrastructure for car, extend infrastructure for rails, or extend infrastructure for bicycles. The urban planner's mix of strategies (choice of investments into different modes) drives the direction of the

[7] For a list of sources see https://en.wikipedia.org/wiki/Modal_share

space allocation dot, which in turn influences the modal share dot due to induced traffic: investment into highways means increased supply of road space, which induces a higher modal share for cars, while investment into protected bicycle lanes increases sustainable bicycle traffic. Similarly, "road diets" and reducing free parking can cause disappearing traffic (Shoup, 2005; Speck, 2013; Weinberger et al., 2008). The impact and effectiveness of such policies could be analyzed visually in a longitudinal, possibly animated, extension of the mobility triangle. The examples of Figure 4 hint towards a non-linear relationship: in the cities of Los Angeles and Rome, where almost all space is allocated to cars, most people are forced to use them, and the discrepancy between space and modal share (the distance between the two dots) is small. In other cities, however, where there is slightly more space for rails or bicycles, like Budapest or Copenhagen, the corresponding forms of transport are much more widely used, leading to a clearly larger discrepancy. This discrepancy suggests that already a small change in infrastructure can provide large positive effects and return on investment—especially considering how inexpensive it is to build bicycle infrastructure.

7.2. Extension of Spatial Scales

A straightforward extension of the platform would be towards different spatial scales, and would give urban planners more options to explore space inequality and transport policies in neighborhoods, districts, areas of interest, agglomerations, or even whole countries. A freely selectable or importable bounding polygon would make the process maximally flexible.

7.3. Scenario Planning for Performance Targets and Estimating Impacts

For the purpose of urban planning, the most hands-on extension would be a feature for scenario planning (Peterson et al., 2003). Scenarios would allow urban planners to play through different transformations, to understand the potential impact of new space-related policies and how concrete performance targets can be reached. For example, many cities struggle with inadequate housing. A scenario planning tool could allow urban planners to understand what happens if a percentage of parking areas were turned into housing. It also could help to answer several questions: How would the narrowing of roads, the addition of protected cycling lanes, or the transformation of city centers into pedestrian zones, change traffic, pollution, walkability, or livability? What percentage of space would need to be converted, and how much investment would that take, to turn a city into the spatial equivalent of Copenhagen? How much sprawl is there (Gervasoni, Bosch, Fenet, & Sturm, 2017)? All the established urban planning goals and indicators could be incorporated, such as climate protection and pollution reduction, adequate and equitable access to housing and transport, health and safety measures like Vision Zero to reduce fatalities from collisions or the increase of daily walking time, open space and agricultural preservation, economic vitality, or transportation system effectiveness.

7.4. Extended Crowdsourced Data Sets

A platform like the one developed is not limited to transport spaces but could become a general visual framework for any kind of urban space inventory where objects are encoded as polygons, collections of line segments, or nodes. For example, the selection of spaces could be extended to amenities, buildings, land use (Fonte et al., 2017), natural areas, waterways, or any of the other built or natural spatial map features of OSM. However, special care has to be taken to fit, and possibly cut, too large elements such as forests, and to determine spatial extensions in case of zero- or one-dimensional map elements. With cutting edge open-source software packages like OSMnx, of which the platform is also making use, downloading, analyzing and visualizing open VGI has become an easy task (Boeing, 2017).

For the purpose of parking and walking space assessment, a desirable data extension would be on-street parking, sidewalks, and other pedestrian spaces. This data could come from VGI systems other than OSM, from commercial platforms, or from city space inventories. The problem with commercial platforms and official city inventories is that data is typically not accessible to the public. More efforts should be invested into establishing transparency laws and open data initiatives like the New York City OpenData project.[8] Extending crowdsourced data with data collected by governmental bodies would have the beneficial side-effect of decreasing data completeness biases. If no up to date information is available, parking data can also be complemented via growth simulations (Chester et al., 2015).

Implementing urban policies should not only follow objective goals, but should first and foremost be satisfactory for the citizens who live in the city. To balance public and private interests, it is therefore important to measure citizen requests and human perceptions of urban spaces. On the one hand, requests can be measured directly through interview surveys, such as in a recent study conducted in Germany by Gesellschaft für Konsumforschung (2017) which showed that 87% of citizens want more space for pedestrian and bicycle traffic. On the other hand, large-scale measurement processes can be designed efficiently by crowdsourcing (Quercia, O'Hare, & Cramer, 2014; Salesses, Schechtner, & Hidalgo, 2013) or by extraction from user-generated social media, allowing for "obtain[ing] citizens' direct feedback for urban planning and as a supplementary decision support tool for ongoing planning processes using contextual emotion information" (Resch, Summa, Zeile, & Strube, 2016, p. 124).

[8] http://opendata.cityofnewyork.us

7.5. From Crowdsourced to Automatized Quantification

While crowdsourced data always suffers from completeness biases, a potentially truly complete approach to assess urban space distribution could come through the recent revolution in automatized machine learning methods (LeCun, Bengio, & Hinton, 2015). In particular, visual methods like image recognition and computer vision have the potential to transform urban planning through the large-scale, automatized identification of urban patterns and dynamics using satellite (Albert, Kaur, & Gonzalez, 2017) or street-level images (Naik, Kominers, Raskar, Glaeser, & Hidalgo, 2017). Such automatized methods come with the need of human-level validation—crowdsourced or in the form of precise inspection (Gössling et al., 2016)—combining automatized with manual approaches. In the example of the platform, image recognition algorithms could identify and measure all different mobility spaces, including sidewalks and on-street parking spots, from satellite images instead of relying on biased user inputs.

7.6. Beyond Space Equality

When comparing allocation of urban mobility space in absolute values, cars naturally take away more space than any other forms due to the high per person area, as discussed in Section 1.1. Therefore, the observed unequal distribution of allocated space is a consequence of both this inherent spatial imbalance and of policy decisions. Given this insight one might ask: Is our definition of equality unfair? Should we accommodate cars and correct for their inherent spatial requirements? If our aim is to measure absolute allocation of space or to raise public awareness of the issue, the answer should be a resounding "No"—the function of urban transport infrastructure is to move people, not arbitrarily large, mostly empty vehicles. However, it could make sense to extend the platform with a rescaling option to disentangle the contributions of inherent space and of policy decisions to understand if a city is favoring cars even above their inherent space requirements. On the other hand, treating cars at the same level as sustainable forms of transport creates a false equivalence and unreasonably accommodates car culture. Besides an array of problems (see Section 8) cars are infinitely more deadly than pedestrians or bicycles, in terms of pollution and as road hazard, and should therefore be reasonably discriminated against in any discussion of equality. Adding this correcting weight to space equality considerations would mean not only going from a) the status quo of prioritizing cars over sustainable mobility to b) a position where sustainable mobility is on an equal footing with cars, but to c) a prioritization of sustainable mobility over cars. Such reversal of prioritization could ultimately lead to a closer measure of fairness that accounts for human life and the way how citizens are treated, i.e. to optimize for livable cities through a human-centric concept of space equity.

8. Further Issues with and Suggested Solutions for Car-centric Cities

Beyond the spatial inefficiency of cars discussed in Section 1.1., there are a number of further sustainability issues, of which we highlight a few here. For a detailed discussion on the related three dimensions of transport injustice—exposure to traffic risks and pollutants, distribution of space, and the valuation of time—see Gössling (2016).

8.1. Pollution

Every year, ambient air pollution causes 3 million deaths worldwide (World Health Organization [WHO], 2017). The largest contributor to such pollutant-related mortalities is road transportation, dominated by cars, causing a large number of PM2.5-related deaths and ozone-related early deaths as an inherent by-product of combustion processes (Molina & Molina, 2004; Caiazzo, Ashok, Waitz, Yim, & Barrett, 2013). A further major problem of vehicular pollution is the rising concentration of greenhouse gases in the atmosphere, contributing to climate change (Moriarty & Honnery, 2008).

For dealing with local health hazards, policy efforts such as a series of European Emission Standards are ongoing to reduce PM emissions of vehicle engines (Piock et al., 2011). Concerning greenhouse gases, a wide mix of technical solutions for improving fuel efficiency in cars has been developed and discussed, including alternative fuels, hydrogen fuel-celled vehicles or hybrid electric vehicles (Moriarty & Honnery, 2008). While fuel efficiency has the potential to be improved by a factor of 4 over the following decades (Åkerman & Höjer, 2006), Moriarty and Honnery (2013) argue that, despite such technical advances, an implementation of "green cars" is not a reasonable solution on a global scale, as several key variables cannot be assumed to be constant. For example, the observed growth of motorization and decrease of car occupancy rates will counteract all technical efforts spent on optimizing fuel efficiency—whether on the tank-to-wheel or well-to-tank level. Therefore, feasible solutions should be non-technical such as transport policy changes to reduce passenger travel levels (Moriarty & Honnery, 2013). However, a recent example in China shows how massive government subsidies and non-monetary incentives can also provide a non-technical solution to boost the explosive adoption of electric vehicles (Wang, Sperling, Tal, & Fang, 2017). In any case, aggressive, visionary policymaking and considerable investments will be needed to achieve success (Fulton, Mason, & Meroux, 2017). How frictionless such policies could be implemented in individualistic Western societies is an open question.

8.2. Road Fatalities

The ninth leading cause of death globally is road traffic crashes, causing 1.25 million people to die every year

(WHO, 2015). Road fatalities are facilitated by a combination of poor safety regulations and inadequate road and vehicle standards, preferentially putting at risk the most vulnerable road users: motorcyclists, pedestrians, and cyclists (WHO, 2015). Recent social psychology research on driver attitudes analyzed the competition for mobility space from the perspective of social dominance theory, further showing that "drivers might view bicyclists as not just a momentary annoyance, but a threat to their social identity as a driver" (Goddard, 2017, p. 17), suggesting a "significant influence on bicycling uptake and bicyclist safety" (Goddard, 2017, p. 152) on top of the inherent physical risks.

In a future where all cars are fully autonomous, the yearly 1.25 million traffic fatalities could become a thing of the past, or could at least be severely reduced (Litman, 2017a). Although such hopes are mostly based on speculation, recent disengagement self-reports from autonomous car testing programs show a trend towards improved technological capabilities (Davies, 2017). However, the fundamental question is: How feasible is the diffusion of self-driving technology globally, and is that preferable to prioritizing infrastructure for the already proven, low-risk forms of mobility? Further, how is technology that was tested in ideal, Western conditions, performing in sub-optimal infrastructure and traffic conditions that exist in most parts of the world (WHO, 2015)? Answers to these questions are yet unclear.

8.3. Lack of Health Benefits

Vehicular mobility does not enjoy the extensive advantages of walking and in particular of cycling which includes fitness benefits and benefits in cardiovascular risk factors (Oja et al., 2011). As Gössling and Choi (2015) have demonstrated via cost-benefit analysis, this is also an economic argument: taking into account health benefits from cycling together with costs on climate change and road crashes, car driving is six times more costly to society (Euro 0.50/km) than cycling (Euro 0.08/km). Although an extra health benefit should not have to be a requirement for designing transport infrastructure, making people sit less still during their travels can be an effective health intervention, especially relevant in today's global obesity epidemic (Frank, Andresen, & Schmid, 2004). There is no obvious solution to the car's lack of health benefits.

8.4. Usage Inefficiency: Sharing and Self-Driving Cars

There are two main ingredients to how usage inefficiency (see Section 1.1) could be overcome. First, focus on increasing occupancy rates through sharing. Although formal ride sharing programs in the US have a long history of failure (Moriarty & Honnery, 2008), strict high-occupancy policies in Jakarta have shown drastic improvements on city-wide traffic (Hanna, Kreindler, & Olken, 2017), again demonstrating the power of non-technical solutions. Apart from such centralized policies, memberships in formal car and ride sharing programs are increasing in Western countries (Shaheen & Cohen, 2007) and are becoming socially acceptable on a large scale through recent advances in information technology and the widespread use of smart phones (Ratti & Biderman, 2017), coupled with changes in behavior that seeks access to mobility instead of ownership of a vehicle (Botsman & Rogers, 2010). The boom of car sharing in the last decade shows that services like car2go, where one shared vehicle can replace up to 11 privately owned vehicles (Martin & Shaheen, 2016), could replace a substantial part of current vehicles in the long run, significantly reducing the need for parking. On the level of taxis, ride sharing scenarios have already been implemented, and their potential benefits have been quantified rigorously—yielding substantial possible improvements in terms of reducing trips and pollution (Santi et al., 2014). Although in individualistic societies consistent success of ride sharing might be difficult to achieve due to locked in user expectations (Epprecht, Von Wirth, Stünzi, & Blumer, 2014), Didi Chuxing self-reported massive savings in China in the order of 510 million liters of fuel over a year (World Economic Forum, 2016).

The second ingredient to better usage efficiency could come in the form of autonomous cars. If all existing vehicles were turned into a self-driving, public fleet, cars would not be bound to a specific human driver anymore, allowing to serve the emerging mobility needs of citizens on the fly, and to even complete freight transport during off-peak times. Combined with the possibility to share, simulated optimal scenarios suggest that the same vehicular mobility needs of today could be delivered with a 70% to 90% smaller fleet (OECD, 2015; Spieser et al., 2014). This more efficient use of cars could mean a reduction of massive amounts of parking space, up to 93% in the case of Stuttgart (Friedrich & Hartl, 2016). At the same time, these scenarios also show that the overall volume of car travel would likely increase due to repositioning and service trips (OECD, 2015). Although public test drives of self-driving cars are being already deployed (Davies, 2017), the acceptance (Epprecht et al., 2014) and large-scale impact on society, with possible substitution and rebound effects, will likely not be clear until considerable numbers are tested in actual use (Ratti & Biderman, 2017).

8.5. No Solution for Spatial Inefficiency

Unfortunately, there is no solution for the issue of spatial inefficiency, making the car the least sustainable form of transportation (Banister, 2005). Although spatial inefficiency might be reduced by sharing and more efficient space use on the road through reduced response times of autonomous vehicles by velocity matching, swarming (Ulbrich, Rotter, & Rojas, 2016), and automatized intersection design (Ratti & Biderman, 2017; Tachet et al., 2016), these efforts would be counteracted by the global increase in motorization, following the same arguments as above (Moriarty & Honnery, 2013). At the same time, increasing the efficiency of cars induces traffic, further reinforcing rather than curbing unsustainable car-centric city planning (Fulton, Mason, & Meroux, 2017). The sprawling space requirements of the car ultimately renders it inferior, even if implemented in the most efficient form, i.e. being environmentally-friendly, public, shared, and autonomous. Thus, for a sustainable urban future, it is paramount to focus on policy solutions and to prioritize the proven methods of mobility: walking, cycling, and mass transit. Such solutions, in particular those aimed at reclaiming space to make cities less car-dependent and more livable, include road diets, abolishing minimum parking requirements and free parking (Chester et al., 2015; Shoup, 2005), or localized car bans if feasible (Speck, 2013) that take into account behavioral responses to be effective (Guerra & Millard-Ball, 2017). Although the high level of car use in cities today is uneconomic and unsustainable, limited numbers of cars might remain useful in special scenarios such as providing mobility for elderly or disabled people (Kamruzzaman, Yigitcanlar, Yang, & Mohamed, 2016). In practice, pricing policies could be easily designed to correct the existing market distortions that currently over-subsidize cars, incorporating congestion, roadway costs, accident risk, parking, pollution, and fuel externalities, towards multimodal and socially optimal transport markets (Gössling, 2015; Litman, 2017b).

9. Conclusion

In this article we explored the spatial implications of almost a century of car-centric urban planning. In particular we focused on how to measure and visualize the inequality of urban mobility spaces on a large scale through crowdsourced data gathered via OSM. We described how the recently developed open-source online platform What the Street!? packs and coils all of a city's mobility spaces of cars, rails, and bicycles, and compares these spaces with each other and with their modal shares, making easily accessible the worldwide privileged allocation of urban space towards cars. The platform also highlights the massive spaces wasted through parking caused by inefficient use of cars. Finally, we analyzed guesses from site visitors, showing that this inequality of space is commonly underestimated. This biased perception of space possibly reinforces conservative, unsustainable transport planning. We showed how open volunteered geographic information enables public engagement with the issue of urban transport inequality, and how it could become a vital part of future planning tools, complementing traditional and automatic methods of urban land use assessment and transport planning.

We have added a spatial perspective to the mounting evidence that car-centric urban planning—which treats the city as a linear machine (Batty, 2013)—is one of the 20th century's most impactful tragedies of the commons come true. It is responsible for a good part of the unfolding global catastrophe of climate change and has created cities with substandard living conditions stuck in gridlock. Initially being an issue of transport inequality that benefited the car industry and a privileged, vehicle-owning segment of the population at the expense of the poor, the paralysis of cities through the inefficiency of vehicular traffic has since long started to hurt all citizens alike. This outcome calls for the benefits of applying the scientific method on an all-encompassing basis instead of relying on myopic transport engineering (Speck, 2013) to make urban transport planning sustainable: 1) gather and analyze data, 2) understand which mix of mobility forms works best for the whole city in the long term, and 3) allocate space and develop infrastructure as needed. Most importantly, this process needs to be agnostic, i.e. without unreasonable prioritization of one form of mobility over others.

Although technical approaches like shared, self-driving cars could free up massive parking spaces, they alone will not make urban transport sustainable. Such technical fixes bring at best temporary benefits, are infeasible to implement globally, and carry the danger of reinforcing the status quo of monocultural, car-centric urban planning and its negative consequences. With the majority of people now living in cities, along increasing worldwide urbanization and motorization, it is high time to reverse the systemic misdevelopments of 20th century urban planning with bold policy making that implements sustainable, society-wide optimal, urban transport systems.

Acknowledgments

I thank moovel and the whole development team of What the Street!?, in particular Stephan Bogner, Benedikt Groß, and Thibault Durand, who were indispensable in the professional implementation of the visualization techniques. I am grateful for collaboration and fruitful discussions on the topic with the moovel lab team, especially with Raphael Reimann, Stephan Bogner, Benedikt Groß, Joey Lee, Tilman Häuser, and Eileen Mandir. I thank Raphael Reimann for providing the guess data set. I thank Thomas Rooney and Roberta Sinatra for helpful feedback on the manuscript, and Lara Szell for the encouragement. Figures 1–5 and 7 are derived from What the Street!? © moovel, MIT License.

Conflict of Interests

The author is a main developer of What the Street!? and was paid employee (researcher in residence) at moovel Group GmbH, subsidiary of Daimler AG, during the creation of the platform. This research was conducted after What the Street!? had been implemented; the research and the decision to publish it was done independently of moovel Group GmbH or Daimler AG.

References

Agentur für clevere Städte. (2014). *Wem gehört die Stadt? Der Flächen-Gerechtigkeits-Report*. Retrieved from https://www.clevere-staedte.de/files/tao/img/blog-news/dokumente/2014-08-05_Flaechen-Gerechtigkeits-Report.pdf

Åkerman, J., & Höjer, M. (2006). How much transport can the climate stand? Sweden on a sustainable path in 2050. *Energy Policy*, *34*(14), 1944–1957. doi:10.1016/j.enpol.2005.02.009

Albert, A., Kaur, J., & Gonzalez, M. (2017). Using convolutional networks and satellite imagery to identify patterns in urban environments at a large scale. *arXiv preprint arXiv:1704.02965*.

Banister, D. (2005). *Unsustainable transport: City transport in the new century*. Abingdon and New York: Routledge.

Bates, J., & Leibling, D. (2012). *Spaced out—Perspectives on parking policy*. London: RAC Foundation. Retrieved from http://www.racfoundation.org/research/mobility/spaced-out-perspectives-on-parking

Batty, M. (2013). *The new science of cities*. Cambridge, MA: MIT Press.

Boeing, G. (2017). OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, *65*, 126–139. doi:10.1016/j.compenvurbsys.2017.05.004

Botsman, R., & Rogers, R. (2010). *What's mine is yours: The rise of collaborative consumption*. New York: HarperCollins.

Caiazzo, F., Ashok, A., Waitz, I. A., Yim, S. H., & Barrett, S. R. (2013). Air pollution and early deaths in the United States. Part I: Quantifying the impact of major sectors in 2005. *Atmospheric Environment*, *79*, 198–208. doi:10.1016/j.atmosenv.2013.05.081

Chester, M., Fraser, A., Matute, J., Flower, C., & Pendyala, R. (2015). Parking infrastructure: A constraint on or opportunity for urban redevelopment? A study of Los Angeles County parking supply and growth. *Journal of the American Planning Association*, *81*(4), 268–286. doi:10.1080/01944363.2015.1092879

Colville-Andersen, M. (2013). The arrogance of space. *The blog by Copenhagenize Design Co*. Retrieved from http://www.copenhagenize.com/2013/03/the-arrogance-of-space.html

Colville-Andersen, M. (2014). The arrogance of space—Paris, Calgary, Tokyo. *The blog by Copenhagenize Design Co*. Retrieved from http://www.copenhagenize.com/2014/09/the-arrogance-of-space-paris-calgary.html

Colville-Andersen, M. (2017). Arrogance of space—Copenhagen—Hans Christian Andersen Boulevard. *The blog by Copenhagenize Design Co*. Retrieved from http://www.copenhagenize.com/2017/05/arrogance-of-space-copenhagen-hans.html

Couper, M. P., Tourangeau, R., Conrad, F. G., & Singer, E. (2006). Evaluating the effectiveness of visual analog scales: A web experiment. *Social Science Computer Review*, *24*(2), 227–245. doi:10.1177/0894439305281503

Davies, A. (2017). The numbers don't lie: Self-driving cars are getting good. *Wired*. Retrieved from https://www.wired.com/2017/02/california-dmv-autonomous-car-disengagement/

Epprecht, N., Von Wirth, T., Stünzi, C., & Blumer, Y. B. (2014). Anticipating transitions beyond the current mobility regimes: How acceptability matters. *Futures*, *60*, 30–40. doi:10.1016/j.futures.2014.04.001

Fonte, C. C., Patriarca, J., Minghini, M., Antoniou, V., See, L., & Brovelli, M. A. (2017). Using OpenStreetMap to create land use and land cover maps: Development of an application. In C. E. C. Campelo, M. Bertolotto, & P. Corcoran (Eds.), *Volunteered geographic information and the future of geospatial data* (pp. 113–137). Hershey, PA: IGI Global.

Frank, L. D., Andresen, M. A., & Schmid, T. L. (2004). Obesity relationships with community design, physical activity, and time spent in cars. *American Journal of Preventive Medicine*, *27*(2), 87–96. doi:10.1016/j.amepre.2004.04.011

Friedrich, M., & Hartl, M. (2016). *MEGAFON—Modellergebnisse geteilter autonomer Fahrzeugflotten des oeffentlichen Nahverkehrs*. University of Stuttgart. Retrieved from http://www.isv.uni-stuttgart.de/vuv/news/Downloads/MEGAFON_Abschlussbericht_V028_20161212.pdf

Fulton, L., Mason, J., & Meroux, D. (2017). *Three revolutions in urban transportation*. Retrieved from https://www.itdp.org/wp-content/uploads/2017/04/ITDP-3R-Report-FINAL.pdf

Gervasoni, L., Bosch, M., Fenet, S., & Sturm, P. (2017). Calculating spatial urban sprawl indices using open data. In *15th International Conference on Computers in Urban Planning and Urban Management*. Adelaide, Australia.

Gesellschaft für Konsumforschung (2017). *Platz für Begegnungen*. Retrieved from https://www.adfc.de/presse/pressemitteilungen/gfk-umfrage–mehr-platz-fuer-begegnungen-weniger-fuer-autos

Glaeser, E. L., & Kahn, M. E. (2004). Sprawl and urban growth. In J. V. Henderson & J. F. Thisse (Eds.), *Handbook of regional and urban economics* (pp. 2481–2527). doi:10.1016/S1574-0080(04)80013-0

Goddard, T. B. (2017). *Drivers' attitudes and behaviors toward bicyclists: Intermodal interactions and implications for road safety* (Ph.D. dissertation), Transportation Research and Education Center, Portland, Oregon, USA.

Gössling, S. (2016). Urban transport justice. *Journal of Transport Geography*, *54*, 1–9. doi:10.1016/j.jtrangeo.2016.05.002

Gössling, S., & Choi, A. S. (2015). Transport transitions in Copenhagen: Comparing the cost of cars and bicycles. *Ecological Economics*, *113*, 106–113.

doi:10.1016/j.ecolecon.2015.03.006

Gössling, S., Schröder, M., Späth, P., & Freytag, T. (2016). Urban space distribution and sustainable transport. *Transport Reviews*, *36*(5), 659–679. doi:10.1080/01441647.2016.1147101

Guerra, E., & Millard-Ball, A. (2017). Getting around a license-plate ban: Behavioral responses to Mexico City's driving restriction. *Transportation Research Part D: Transport and Environment*, *55*, 113–126. doi:10.1016/j.trd.2017.06.027

Haklay, M. (2010). How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and Planning B: Planning and Design*, *37*(4), 682–703. doi:10.1068/b35097

Haklay, M., Basiouka, S., Antoniou, V., & Ather, A. (2010). How many volunteers does it take to map an area well? The validity of Linus' law to volunteered geographic information. *The Cartographic Journal*, *47*(4), 315–322. doi:10.1179/000870410X12911304958827

Hanna, R., Kreindler, G., & Olken, B. A. (2017). Citywide effects of high-occupancy vehicle restrictions: Evidence from "three-in-one" in Jakarta. *Science*, *357*(6346), 89–93. doi:10.1126/science.aan2747

Hutton, B. (2013). *Planning sustainable transport*. Abingdon and New York: Routledge.

Kamruzzaman, M., Yigitcanlar, T., Yang, J., & Mohamed, M. A. (2016). Measures of transport-related social exclusion: A critical review of the literature. *Sustainability*, *8*(7), 696. doi:10.3390/su8070696

Kondor, D., Zhang, H., Tachet, R., Santi, P., & Ratti, C. (2017). Estimating savings in parking demand using shared vehicles for home-work commuting. *arXiv:1710.04983v2*.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. doi:10.1038/nature14539

Litman, T. (2017a). *Autonomous vehicle implementation predictions*. Victoria, BC: Victoria Transport Policy Institute. Retrieved from http://www.vtpi.org/avip.pdf

Litman, T. (2017b). *Socially optimal transport prices and markets*. Victoria, BC: Victoria Transport Policy Institute. Retrieved from http://www.vtpi.org/sotpm.pdf

López-Camacho, E., Ochoa, G., Terashima-Marín, H., & Burke, E. K. (2013). An effective heuristic for the two-dimensional irregular bin packing problem. *Annals of Operations Research*, *206*(1), 241–264. doi:10.1007/s10479-013-1341-4

Louf, R., & Barthelemy, M. (2013). Modeling the polycentric transition of cities. *Physical Review Letters*, *111*(19), 198702. doi:10.1103/PhysRevLett.111.198702

Martin, E., & Shaheen, S. (2016). *Impacts of car2go on vehicle ownership, modal shift, vehicle miles traveled, and greenhouse gas emissions: An analysis of five North American cities* (Working Paper). Berkeley, CA: Transportation Sustainability Research Center.

Martin, G. (2007). Global motorization, social ecology and China. *Area*, *39*(1), 66–73. doi:10.1111/j.1475-4762.2007.00718.x

Mohl, R. A. (2002). *The interstates and the cities: Highways, housing and the freeway revolt* (Research Report). Washington, DC: Poverty and Race Research Action Council. Retrieved from http://www.prrac.org/pdf/mohl.pdf

Molina, M. J., & Molina, L. T. (2004). Megacities and atmospheric pollution. *Journal of the Air & Waste Management Association*, *54*(6), 644–680. doi:10.1080/10473289.2004.10470936

Moriarty, P., & Honnery, D. (2008). The prospects for global green car mobility. *Journal of Cleaner Production*, *16*(16), 1717–1726. doi:10.1016/j.jclepro.2007.10.025

Moriarty, P., & Honnery, D. (2013). Greening passenger transport: A review. *Journal of Cleaner Production*, *54*, 14–22. doi:10.1016/j.jclepro.2013.04.008

Naik, N., Kominers, S. D., Raskar, R., Glaeser, E. L., & Hidalgo, C. A. (2017). Computer vision uncovers predictors of physical urban change. *Proceedings of the National Academy of Sciences*, *114*(29), 7571–7576. doi:10.1073/pnas.1619003114

National Association of City Transportation Officials. (2017). *Designing for all ages & abilities*. New York: NACTO. Retrieved from https://nacto.org/2017/12/06/designing-for-all-ages-and-abilities/

Norton, P. D. (2007). Street rivals: Jaywalking and the invention of the motor age street. *Technology and Culture*, *48*(2), 331–359.

Organisation for Economic Co-operation and Development. (2015). *Urban mobility system upgrade—How shared self-driving cars could change city traffic*. Paris: OECD. Retrieved from https://www.itf-oecd.org/sites/default/files/docs/15cpb_self-drivingcars.pdf

Oja, P., Titze, S., Bauman, A., de Geus, B., Krenn, P., Reger-Nash, B., & Kohlberger, T. (2011). Health benefits of cycling: A systematic review. *Scandinavian Journal of Medicine & Science in Sports*, *21*(4), 496–509. doi:10.1111/j.1600-0838.2011.01299.x

Peterson, G. D., Cumming, G. S., & Carpenter, S. R. (2003). Scenario planning: A tool for conservation in an uncertain world. *Conservation Biology*, *17*(2), 358–366. doi:10.1046/j.1523-1739.2003.01491.x

Piock, W., Hoffmann, G., Berndorfer, A., Salemi, P., & Fusshoeller, B. (2011). Strategies towards meeting future particulate matter emission requirements in homogeneous gasoline direct injection engines. *SAE International Journal of Engines*, *4*(1), 1455–1468. doi:10.4271/2011-01-1212.

Pucher, J., & Buehler, R. (2016). Safer cycling through improved infrastructure. *American Journal of Public Health*, *106*(12), 2089–2091. doi:10.2105/AJPH.2016.303507

Quattrone, G., Capra, L., & De Meo, P. (2015). There's no such thing as the perfect map: Quantifying bias in spatial crowd-sourcing datasets. *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*, (pp. 1021–1032). New York: ACM.

Quercia, D., O'Hare, N. K., & Cramer, H. (2014). Aesthetic capital: What makes London look beautiful, quiet, and happy? In *Proceedings of the 17th ACM conference on computer supported cooperative work & social computing*, (pp. 945–955). New York: ACM.

Ratti, C., & Biderman, A. (2017). From parking lot to paradise. *Scientific American*, *317*(1), 54–59. doi:10.1038/scientificamerican0717-54

Resch, B., Summa, A., Zeile, P., & Strube, M. (2016). Citizen-centric urban planning through extracting emotion information from Twitter in an interdisciplinary space-time-linguistics algorithm. *Urban Planning*, *1*(2), 114–127. doi:10.17645/up.v1i2.617

Reynolds, C. C., Harris, M. A., Teschke, K., Cripton, P. A., & Winters, M. (2009). The impact of transportation infrastructure on bicycling injuries and crashes: A review of the literature. *Environmental Health*, *8*(1), 47. doi:10.1186/1476-069X-8-47

Salesses, P., Schechtner, K., & Hidalgo, C. A. (2013). The collaborative image of the city: Mapping the inequality of urban perception. *PloS one*, *8*(7), e68400. doi:10.1371/journal.pone.0068400

Santi, P., Resta, G., Szell, M., Sobolevsky, S., Strogatz, S. H., & Ratti, C. (2014). Quantifying the benefits of vehicle pooling with shareability networks. *Proceedings of the National Academy of Sciences*, *111*(37), 13290–13294. doi:10.1073/pnas.1403657111

Senaratne, H., Mobasheri, A., Ali, A. L., Capineri, C., & Haklay, M. (2017). A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*, *31*(1), 139–167. doi:10.1080/13658816.2016.1189556

Shaheen, S., & Cohen, A. (2007). Growth in worldwide carsharing: An international comparison. *Transportation Research Record: Journal of the Transportation Research Board*, (1992), 81–89. doi:10.3141/1992-10

Shoup, D. C. (2005). *The high cost of free parking*. Chicago: Planners Press.

Speck, J. (2013). *Walkable city: How downtown can save America, one step at a time*. Basingstoke, UK: Macmillan.

Spieser, K., Treleaven, K., Zhang, R., Frazzoli, E., Morton, D., & Pavone, M. (2014). Toward a systematic approach to the design and evaluation of automated mobility-on-demand systems: A case study in Singapore. In G. Meyer & S. Beiker (Eds.), *Road vehicle automation* (pp. 229–245). Cham: Springer.

Tachet, R., Santi, P., Sobolevsky, S., Reyes-Castro, L. I., Frazzoli, E., Helbing, D., & Ratti, C. (2016). Revisiting street intersections using slot-based systems. *PloS one*, *11*(3), e0149607. doi:10.1371/journal.pone.0149607

Teschke, K., Harris, M. A., Reynolds, C. C., Winters, M., Babul, S., Chipman, M., . . . Cripton, P. A. (2012). Route infrastructure and the risk of injuries to bicyclists: a case-crossover study. *American Journal of Public Health*, *102*, 2336–2343.

Ulbrich, F., Rotter, S. S., & Rojas, R. (2016). Adapting to the traffic swarm: Swarm behaviour for autonomous cars. In Y. Tan (Ed.), *Handbook of research on design, control, and modeling of swarm robotics* (pp. 263–285). Hershey, PA: IGI Global.

Urry, J. (2013). *Societies beyond oil: Oil dregs and social futures*. London: Zed Books.

Wainer, H. (2005). *Graphic discovery: A trout in the milk and other visual adventures*. Princeton University Press.

Wang, Y., Sperling, D., Tal, G., & Fang, H. (2017). China's electric car surge. *Energy Policy*, *102*, 486–490. doi:10.1016/j.enpol.2016.12.034

Weigele, S. (2014) . *Urbane Mobilität im Umbruch?* Hamburg: civity Management Consultants. Retrieved from http://www.civity.de/de/publikationen/2014/08/matters_no_1_-_urbane_mobilit%C3%A4t_im_umbruch

Weinberger, R., Seaman, M., Johnson, C., & Kaehny, J. (2008). *Guaranteed parking-Guaranteed driving*. New York: Transportation Alternatives.

Winters, M., Babul, S., Becker, H. J. E. H., Brubacher, J. R., Chipman, M., Cripton, P., . . . Monro, M. (2012). Safe cycling: how do risk perceptions compare with observed risk. *Canadian Journal of Public Health*, *103*(9), S42–7.

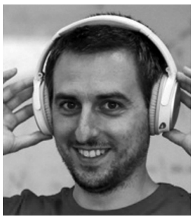
World Economic Forum. (2016). *Understanding the sharing economy*. Cologny, Switzerland: World Economic Forum.

World Health Organization. (2015). *Global status report on road safety 2015*. Geneva, Switzerland: WHO.

World Health Organization. (2017). *World health statistics 2017*. Geneva, Switzerland: WHO.

Zhang, H., & Malczewski, J. (2017). Quality evaluation of volunteered geographic information: The case of OpenStreetMap. In In C. E. C. Campelo, M. Bertolotto, & P. Corcoran (Eds.), *Volunteered geographic information and the future of geospatial data* (pp. 19–46). Hershey, PA: IGI Global.

About the Author



Michael Szell is Assistant Professor at Central European University (Center for Network Science), Adjunct Visiting Professor at Northeastern University (Center for Complex Network Research), External Faculty at the Complexity Science Hub Vienna, and Researcher at the Hungarian Academy of Sciences. He was part of the development team of What the Street!? (whatthestreet.moovellab.com) as researcher in residence at moovel lab. His research focuses on quantifying human behavior and social networks, urban sustainability and mobility, with a particular focus on data visualization.

Urban Planning (ISSN: 2183–7635)
2018, Volume 3, Issue 1, Pages 21–33
DOI: 10.17645/up.v3i1.1231

Article

Investigating the Emotional Responses of Individuals to Urban Green Space Using Twitter Data: A Critical Comparison of Three Different Methods of Sentiment Analysis

Helen Roberts [1,*], Bernd Resch [2,3], Jon Sadler [1], Lee Chapman [1], Andreas Petutschnig [2] and Stefan Zimmer [2]

[1] School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, B15 2TT, UK;
E-Mails: hxr133@bham.ac.uk (H.R.), j.p.sadler@bham.ac.uk (J.S.), l.chapman@bham.ac.uk (L.C.)
[2] Department of Geoinformatics, University of Salzburg, 5020 Salzburg, Austria; E-Mails: bernd.resch@sbg.ac.at (B.R.), andreas.petutschnig@sbg.ac.at (A.P.), stefan.zimmer@sbg.ac.at (S.Z.)
[3] Center for Geographic Analysis, Harvard University, Cambridge, MA 02138, USA; E-Mail: bresch@fas.harvard.edu

* Corresponding author

Submitted: 19 October 2017 | Accepted: 22 January 2018 | Published: 29 March 2018

Abstract
In urban research, Twitter data have the potential to provide additional information about urban citizens, their activities, mobility patterns and emotion. Extracting the sentiment present in tweets is increasingly recognised as a valuable approach to gathering information on the mood, opinion and emotional responses of individuals in a variety of contexts. This article evaluates the potential of deriving emotional responses of individuals while they experience and interact with urban green space. A corpus of over 10,000 tweets relating to 60 urban green spaces in Birmingham, United Kingdom was analysed for positivity, negativity and specific emotions, using manual, semi-automated and automated methods of sentiment analysis and the outputs of each method compared. Similar numbers of tweets were annotated as positive/neutral/negative by all three methods; however, inter-method consistency in tweet assignment between the methods was low. A comparison of all three methods on the same corpus of tweets, using character emojis as an additional quality control, identifies a number of limitations associated with each approach. The results presented have implications for urban planners in terms of the choices available to identify and analyse the sentiment present in tweets, and the importance of choosing the most appropriate method. Future attempts to develop more reliable and accurate algorithms of sentiment analysis are needed and should focus on semi-automated methods.

Keywords
emotions; sentiment analysis; Twitter; urban green space; urban planning

Issue
This article is part of the issue "Crowdsourced Data and Social Media for Participatory Urban Planning", edited by Bernd Resch (University of Salzburg, Austria), Peter Zeile (Karlsruhe Institute of Technology, Germany) and Ourania Kounadi (University of Salzburg, Austria).

1. Introduction

1.1. Twitter, Sentiment Analysis and Urban Green Space

Sentiment analysis describes the field of study concerned with analysing the opinions, attitudes and emotions of individuals towards entities such as products, services, organisations, locations and events (Liu, 2012). Over the last two decades, the field has become increasingly active given the vast real-world applications to a plethora of disciplines, such as politics, economics, business, healthcare and urban planning. Increased engagement with sentiment analysis has also coincided with the rapid growth in social networks, without which a lot of the recent research would not have been possible. For the first time in human history researchers have access to huge volumes of freely accessible data published by individuals online.

The increase in social media sites such as Twitter has led to the internet becoming a place of increased expression and opinion sharing on a vast range of topics (Pak & Paroubek, 2010). This phenomenon is providing new sources of text which can be used to gauge public opinion through sentiment analysis (Zhang, Riddhiman, Dekhil, Hsu, & Liu, 2011). Recent studies have indicated the potential and versatility of tweets in examining emotions. These include: a variety of economic (Bollen, Mao, & Zeng, 2011; Chung & Liu, 2011; Jansen, Zhang, Sobel, & Chowdury, 2009) and social (Thelwall, 2014) contexts, examining emotional responses to specific events, such as political elections (Bruns & Burgess, 2011; Tumasjan, Sprenger, Sandner, & Welpe, 2010; Wang, Can, Kazemzadeh, Bar, & Narayanan, 2012), natural disasters (Mandel et al., 2012; Shalunts, Backfried, & Prinz, 2014) and terrorism events (Cheong & Lee, 2011); and exploring new ways to measure happiness (Dodds, Harris, Kloumann, Bliss, & Danforth, 2011; Mitchell, Frank, Harris, Dodds, & Danforth, 2013; Quercia, Ellis, Capra, & Crowcroft, 2012). Recent research by Roberts, Sadler and Chapman (in press) identified how Twitter data can be successfully used to identify both emotions in tweets; and the cause of these emotions, in relation to green space experience. Following the success of this work, this study investigates the use of three different methods of sentiment analysis in this context. In doing so, different methodologies are explored and their limitations discussed.

The information made available by individuals in their tweets has the potential to provide insights into how urban landscapes are perceived by individuals as they navigate them. The urban landscape is being experienced by an increasing number of individuals as global urban populations continue to expand (UN Habitat, 2016), leading some to question the long-term sustainability of cities (Grimm, Grove, Pickett, & Redman, 2000). Understanding how individuals are responding and relating to city landscapes is a key element for facilitating their design, implementation and management. Urban green spaces in cities provide the opportunity for individuals to have contact with the natural environment (Daniel et al., 2012), a fundamental influence on human well-being (Fuller & Gaston, 2009; Kellert & Wilson, 1995; Wilson, 1984), while the benefits associated with nature and green spaces are a vital component of the ecosystem services they provide to human populations (Costanza et al., 1997; Daily, 1997; Ehrlich & Ehrlich, 1981; MEA, 2005). Despite broad agreement that these cultural ecosystem services are beneficial to urban dwellers (World Health Organisation, 2017) there remains limited methodological progress in capturing the transfer and receipt of these services to populations, largely due to their intangible nature and difficulty in assigning economic value to the benefits they provide (Daniel et al., 2012; Milcu, Hanspach, Abson, & Fischer, 2013). Studies have only recently emerged that consider the effect of number and duration of encounters on ecosystem service receipt (Shanahan, Fuller, Bush, Lin, & Gaston, 2015; Shanahan, Lin, Gaston, Bush, & Fuller, 2014), and at present they remain small scale and highly contextualised. Twitter data have the potential to offer a wider spatial and temporal lens through which responses of people to urban green spaces can be captured.

While environmental cues have a significant impact on how individuals respond to and experience space (Ulrich, 1983), a wide range of other factors are also influential, including weather conditions, group dynamics, types of activities and what people observe happening around them. These factors are hard to study successfully due to limitations on experiment size and cohort selection, so capturing their high spatial and temporal variability has proved challenging (Cohen et al., 2009). As a result, studies lack explorations of the emotional responses of people to urban green spaces and the range of sentiments they can elicit in individuals. Twitter data offers the potential to overcome these limitations and can provide information about how individuals feel while experiencing urban green spaces. The information provided in tweets also has the potential to contextualise why an individual may be experiencing certain emotions and what activities they are engaging in that result in the given response. Such information has significant utility for urban planning. For example, data which provides evidence for the beneficial effects of urban green spaces for urban dwellers can be used to justify their continued presence in the urban landscape amidst intense development pressures. Furthermore, the successful identification of the causes of positive and negative emotions experienced by users of urban green space using Twitter data (Roberts et al., in press), could be used to develop an evidence base from the which planners can create and manage green spaces to promote positive emotional experiences and minimise and remove features which cause negative responses.

Despite the benefits Twitter data offers to researchers, sentiment analysis studies obtained from tweets are not common, especially in an urban context. Nonetheless, studies have utilised tweet text to investigate how public mood varies both spatially (Bertrand, Bialik, Virdee, Gros, & Bar-Yam, 2013) and temporally (Martinez & González, 2013) in urban areas, and to compare how the positivity of Twitter posts by urban citizens varies between different cities (Hollander et al., 2016). Others have used Twitter data alongside additional sources (such as biosensors) to assess how individuals perceive and emotionally respond to cities (Resch, Summa, Zeile, & Strube, 2016), in order to develop more citizen centric approaches to urban planning. For tweets to be a useful source of emotional data to urban planners, methods of sentiment analysis are required which enable the fast, accurate and replicable annotation of tweets.

1.2. Methods of Sentiment Analysis

The possibility of accurately extracting emotions from tweets has been demonstrated in recent studies (e.g., Roberts, Roach, Johnson, Guthrie, & Harabagiu, 2012), which have classified tweets according to a range of readily identifiable and distinct emotions. However, working with such an informal text genre presents new challenges for language processing as the language used by the twitter community is often informal with creative punctuation and spelling, slang, abbreviations and URLs (Rosenthal, Ritter, Nakov, & Stoyanov, 2014). The use of emoticons/emojis also provides an additional challenge for analysts as the emotions they convey can be highly subjective and often context dependent. Debate on how to develop methods which address these challenges and capture the fullest range of responses possible, and how best to mine people's opinions and sentiments is an increasing body of literature.

To compensate for the range of challenges inherent in using Twitter data, approaches to identifying sentiment and emotion are varied, but can broadly be placed into three commonplace methodologies. Firstly, manual annotation requires human annotators to categorise tweets into emotion categories (Jansen et al., 2009; Roberts et al., in press). Fully automated annotation can also be undertaken, relying on algorithms and rules to annotate the emotion in tweets. Many different approaches to fully automated annotation exist, but methods typically rely on n-gram analysis (Barbosa & Feng, 2010) to annotate the emotion in a tweet. Significant limitations have been identified with using both manual and automated sentiment analysis on tweets (and are discussed in detail in subsequent sections). As a result, novel methodologies are being developed to progress tweet sentiment analysis. This study presents one such method, drawing on semi-supervised or machine learning annotation. There are a number of machine learning techniques which can be employed to annotate tweets including Naïve Bayes classification (Go, Bhayani, & Huang, 2009; Pak & Paroubek, 2010), maximum entropy classification (Go et al., 2009), graph based propagation algorithms (Resch et al., 2016) and semantic orientation (Turney, 2002). The method presented herein relies on a graph based semi-supervised learning algorithm (Resch et al., 2016) and is described in full in Section 2.5. The variety of approaches undertaken within these three methodological approaches reflects the complexity inherent in the task.

This article uses tweets relating to urban green spaces to evaluate three different sentiment analysis methods, focusing on the variation in sentiment they indicate, in order to facilitate discussion around the limitations and benefits of each approach. However, this article does not attempt to identify the most effective method for tweets. Instead, the aims of this article are twofold:

1) To compare the outcomes of manual, fully automated and semi-supervised learning methods of sentiment analysis on the same corpus of tweets;
2) To evaluate each method in the context of urban green space research.

The three methods of sentiment analysis presented and compared herein have been chosen as each one is derived from one of the three broad methodologies of sentiment analysis: manual, automated and semi-automated. In this way, a comparison can be made between these differing methodologies in the context of urban green space research; and their potential contribution in providing ways for urban planners to engage meaningfully with social media derived data.

2. Methodology

2.1. Case Study Location

The tweets collated for analysis relate to 60 urban green spaces located in Birmingham, United Kingdom (Figure 1). With a population of approximately 1.1 million people (Office for National Statistics, 2014) the 600 public parks, open spaces and nature reserves within the Birmingham metropolitan area (Birmingham City Council, 2016) provide an important resource for urban citizens in terms of their contribution to cultural ecosystem service provision.

The 60 green spaces were chosen to reflect the diversity of spaces found across the city in terms of their size, habitat type, available facilities and amenities and locations within different types of neighbourhoods. Alongside 46 parks, 14 linear green features were also included for investigation consisting of the footpaths along 4 rivers and 7 canals and 3 cycle ways.

2.2. Tweet Corpus Creation

The tweets used in this study were obtained via Twitters publically accessible REST API. The REST API provides access to a 1% sample of tweets published by users with public profiles, and allows queries to be used to search for specific tweets. Searches made using the REST API are based on relevance and therefore this source of tweets was most appropriate for use in this article. To create the tweet corpus used in this study, English language tweets were downloaded every 10 days from the REST API. During preparation for the tweet data collection various different time scales were used to collect tweets to ascertain the most effective frequency for harvesting tweets. Tests were carried out over a three month trial period to look at which frequency worked best to harvest tweets in terms of minimising duplications and ensuring sufficient capture of the available tweets. Frequencies of 3, 5, 7 and 10 days were tested. This showed that using frequencies of 3, 5 and 7 days were too frequent and re-

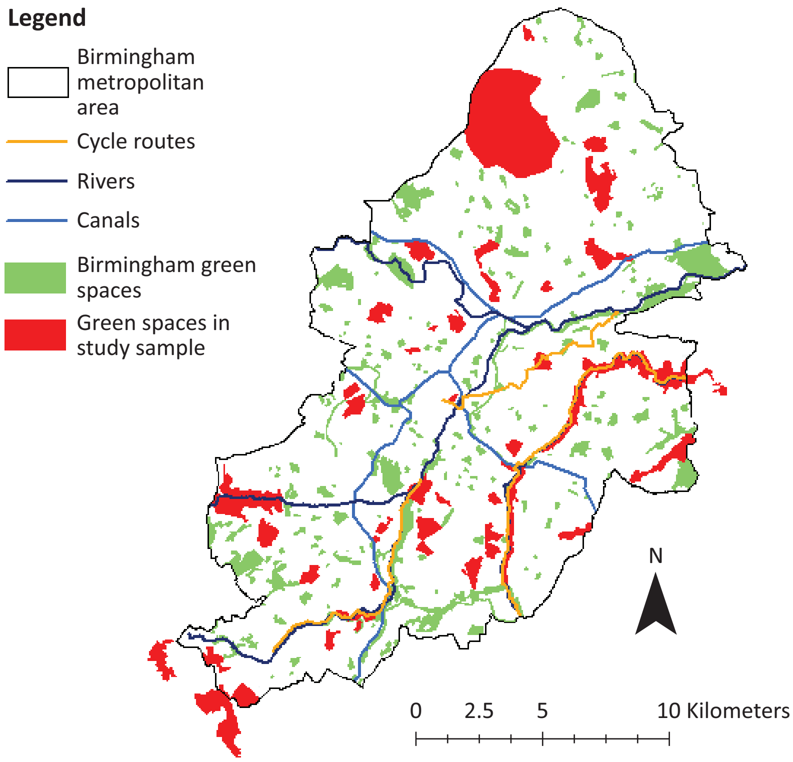


Figure 1. The locations of the green space study sites included in this article.

sulted in large duplications and made unnecessary pre-processing work to remove the duplications. Using the 10 day frequency, there was no lack of tweets compared to searching every 7 days, and given the benefits of this frequency in harvesting the tweets this frequency was used throughout the subsequent data collection period ensuring maximal temporal coverage over a period of 12 months, from June 2015 to May 2016. A search query was used to ensure that the tweets downloaded related to one of the 60 sites included in the study. Therefore, each tweet in the corpus contains specific reference to one of the sixty green spaces included in the sample. Any duplicated tweets were removed during pre-processing. In this way, a corpus of 10268 tweets was generated for use in this study.

2.3. Manual Annotation

During manual annotation, tweets were first assigned into one of three categories: positive, negative or neutral. This annotation was based on the presence of emotive words, emoticons/emojis or meaning. Subsequently, the positive and negative tweets were categorised into distinct emotions. The higher level emotions chosen included five of Ekman's basic emotions (anger, disgust, fear, sadness and happiness (Ekman, 1999; Ekman & Friesen, 1971)), in line with previous research using Twitter data (Roberts et al., 2012; Resch et al., 2016). These emotions are arranged into the ontology shown in Figure 2. In this study, beauty was included an additional sub-category to the positive tweets but outside of the emotions to account for the large amount of tweets referencing the beauty of nature and the landscape (as to be expected for green space). Each tweet could only be assigned into one of these emotion categories based on the strongest present emotion.

Five annotators were used to annotate a random sample of 1,000 tweets, in order to ensure there was sufficient agreement between different annotators in how tweets were categorised. A metric of comparison was

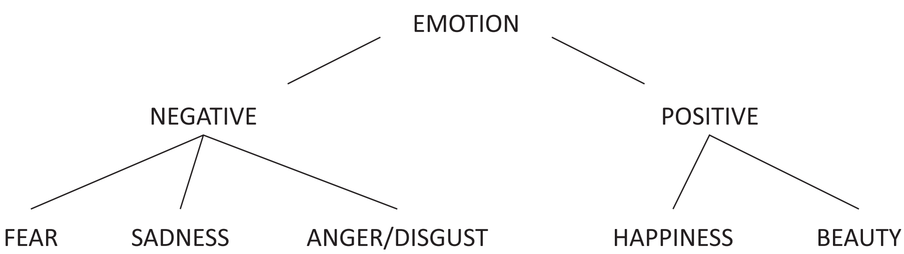


Figure 2. High level emotion ontology for the emotions used in manual and semi-automated tweet annotation.

derived (K = 0.666) suggesting sufficient agreement to assume inter-annotator reliability (Landis & Koch, 1977). Given the identification of sufficient inter-annotator reliability between annotators, and the time required for the task, the remaining tweets were annotated by one annotator. To the authors' knowledge this is largest manually annotated dataset of sentiment present in tweets, providing a robust test set against which other methods can be compared.

2.4. Fully Automated Annotation

For the automated method, an Affective Norms for English Words (ANEW) resource was used as the basis for emotion annotation. The ANEW resource utilised here derives from Warriner, Kuperman and Brysbaert (2013) in which over 13,000 English lemmas were assigned valence scores. Using an automated process these scores were used to annotate the valency of each tweet in the corpus. After assigning each word in each tweet with a valence score, an average valence score was created for each tweet based on the number of words present. Thresholds were then used to place the tweets into positive, neutral and negative categories. Following the thresholds used by Warriner et al. (2013) tweets with scores of ≥6.0 were categorised as positive, scores between 5.9 and 4.9 were categorised as neutral and scores of ≤4.9 were categorised as being negative. Given there remains no robust way to determine specific emotions from numeric scores, this method only annotated the tweets in terms of their positivity as opposed to annotating each with a discrete emotion. The implications of this are discussed in greater detail further on.

2.5. Graph Based Semi-Supervised Learning Annotation

In this method (Resch et al., 2016), a sample of manually annotated tweets was used to train a graph based semi-supervised learning algorithm which annotated the remaining tweets. A sample of 1,000 tweets from the corpus, known as the gold standard, were annotated manually (as described in Section 2.3) and used to train and evaluate the annotation algorithm. This was done to compromise between manual and automated analysis and capture the benefits of each, namely the accuracy of manual annotation and the quickness of automated annotation.

In order to classify tweets according to the emotions they contain a similarity computation was first undertaken, where similarity is defined as the likelihood that two tweets contain the same emotion. The similarity computation comprises three dimensions; linguistic similarity (defined through proven emotion emotion-related linguistic features such as co-occurring words, part-of-speech tags, punctuation, spelling, emojis and n-grams), spatial similarity and temporal similarity (defined through spatial and temporal decay functions according to recent literature). It should be noted that the results presented in this article only used the linguistic feature groups because not all tweets were geolocated, thus lacked the necessary spatial information.

Once the similarity between tweets has been computed, the graph, which creates the input for the semi-supervised learning approach is constructed and is defined by the tweets (nodes) and pairwise similarity values (weighted edges). Assigning emotions to the tweets was undertaken by applying the graph-based semi-supervised learning algorithm Modified Absorption (MAD) using a subset of the gold standard (training dataset) as this method is found to be most effective for graphs where nodes connect to many other nodes (Talukdar & Pereira, 2010). The assigned emotions were then validated using the rest of the gold standard (test dataset) through computing statistical measures including precision, recall, f-measure and micro average precision. The results prove to be better than random and majority baselines which in the understanding of the field of computational linguistics, demonstrates that the methods works well. The developed algorithm outperforming the majority baseline is considered assuring, whereas the better performance compared to random baseline provides strong evidence that the method works well because it demonstrates that the results are not produced by chance, but that significant similarities have been found between pairs of tweets.

Once each tweet had been assigned a discrete emotion using this method, it was then possible to reverse the process and place the tweets into positive, neutral and negative categories using the same ontology as shown in the manual method.

2.6. Analysis

Following presentation of the relevant descriptive statistics for each method, various statistical tests were undertaken to assess the significance of any differences in the assignment of the number of positive, neutral and negative tweets by each of the three methods. Fleiss and Cohen Kappa Indexes were then generated to assess inter-method reliability of tweet assignment into each category between the three methods alongside percentage agreement assessments of the three methods in their annotation of each individual tweet.

3. Results

3.1. Assignment of the Tweets into Positive, Neutral and Negative Categories

Variation existed in the numbers of tweets assigned to into the 'positive', 'neutral' and 'negative' categories by each of the methods (Figure 3). Although for all three methods, the majority of tweets were placed into the 'neutral' category, categorisation of tweets into 'positive' and 'negative' categories showed to be more variable between the three methods (Table 1).

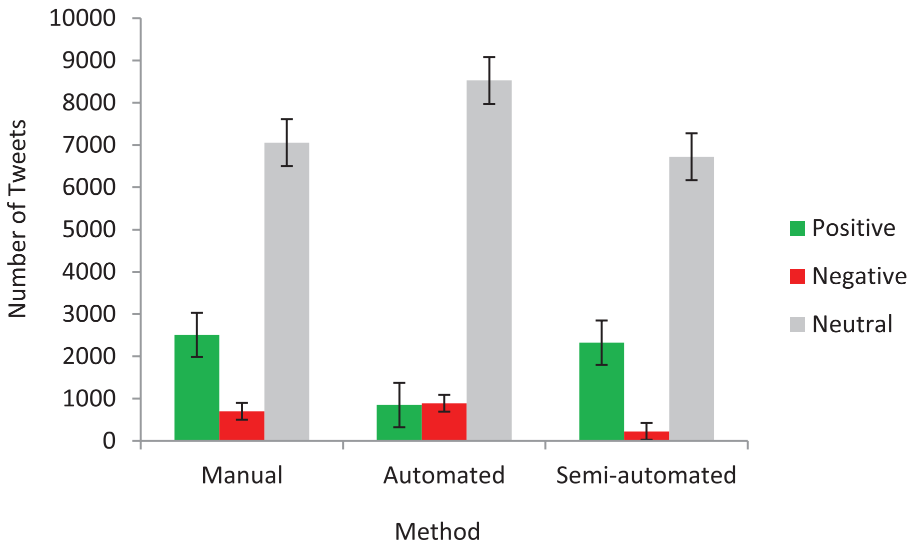


Figure 3. The number of tweets assigned by each method into positive, neutral and negative categories with standard error bars displayed. N (Number of tweets analysed) = 10268, for all methods.

Table 1. The percentage (%) of tweets assigned by each method to positive, neutral and negative categories.

	Manual	Automated	Semi-Automated
Positive	24.4	8.2	25.1
Neutral	68.8	83.0	72.5
Negative	6.8	8.8	2.4

Given that all three methods show some similarity in the numbers of tweets assigned to each category (Figure 3), statistical analysis was undertaken to investigate the significance of the differences identified between the three methods of classification for all three classes: 'positive', 'negative' and 'neutral'. Given that the assumption of homogeneity of variance was not met by the 'positive' datasets, a Welch ANOVA test was used and identified significant difference in the number of tweets annotated as positive by each of the three methods ($F(2,17.867)=39.343$, $p < 0.001$). Post hoc Tukey analysis identified specific significant differences between manual and automated analysis ($p < 0.001$) and automated and semi-automated analysis ($p = 0.001$). There was no significant difference in the number of tweets annotated as 'positive' by the manual and semi-automated methods ($p = 0.76$). Using a one-way ANOVA, no significant differences were identified between the number of tweets classified as being 'neutral' by each method ($F(2,33)=3.216$, $p = 0.053$). Finally, a Kruskal-Wallis H test, given the violated assumption of normality, identified significant differences between the number of tweets classified as 'negative' by the three methods ($\chi^2(2)=16.176$, $p < 0.001$) . These were largest between the automated and semi-automated annotations of negativity.

By making adjustment to the thresholds (Table 2) used to assign the automated tweet scores into the 'positive', 'neutral' and 'negative' categories, it was possible to generate very similar outputs for the manual and fully automated methods (Figure 4), and identify no significant differences in the number of tweets each method assigned to each category.

3.2. Inter-Method Reliability

Consideration of inter-method reliability however, shows a more complex picture. A Fleiss Kappa Index identified very little inter-method agreement ($k = 0.0445$) between the three methods, highlighting that the annotation of each individual tweet into the three different categories by each method differed substantially. Indeed,

Table 2. Original and adjusted thresholds used to assign automated tweet scores into positive, neutral and negative categories.

	Original threshold adapted from Warriner et al. (2013)	Adjusted threshold
Positive assigned tweets	≥6.0	≥5.73
Neutral assigned tweets	≥5.0	≥4.931
Negative assigned tweets	≤4.99	≤4.93

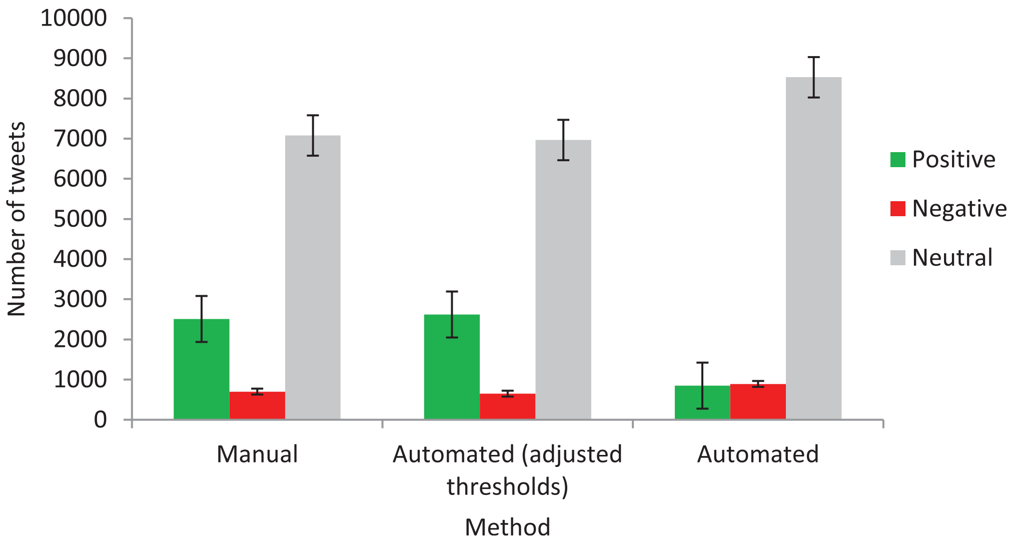


Figure 4. Comparisons of the numbers of tweets assigned to positive, neutral and negative categories by the manual and automated methods using two different thresholds.

only 44.5% of tweets were found to have been assigned the same category by all three methods, with 5.5% of tweets being assigned different categories by all three methods, indicating wide misallocation.

The relatively high percentage agreement compared to the low Fleiss Kappa Index is due to a large number of tweets being annotated as neutral by all three methods. Indeed, further investigation of the 44.5% of tweets which were annotated the same by all three methods revealed the vast majority to have been assigned to the 'neutral' category (98.1%). However, annotations of positive and negative tweets were less similar, suggesting that where emotions were present, the methods showed more variance in identifying them, either annotating them as neutral or with the incorrect polarity of positivity. Positive and negative annotation agreement between all three methods was extremely low at 1.9% and 0% respectively.

Interestingly, the low percentage in the agreement of tweets remained following the adjustment of the automated thresholds. The adjusted threshold annotations showed most similarity with the manual annotations. Again, however, only 56.8% of tweets were placed in the same category by both methods; showing that despite increasing similarity in number of tweets assigned to each category by each method, altering the thresholds used to assign tweets into 'positive', 'neutral' and 'negative' categories had no effect on increasing the percentage agreement of tweet assignment between the manual and fully automated methods.

Cohen Kappa tests were undertaken to see if the inter-method reliability was higher between any two specified annotation methods. The highest inter-method reliability was found to be between the manual and semi-automated methods (K = 0.136), compared to similarity between manual and automated (K = 0.0814), and semi-automated and automated methods (K = –0.00784). However, all these Kappa Indices are low (McHugh, 2012) and there remains large variation in the way each method assigns individual tweets into 'positive', 'neutral' or 'negative' categories, despite the appearance of similarity in Figure 3.

3.3. Quality Control Using Character Emojis

By way of a quality control measure, assessment was undertaken on just the tweets containing objective character emojis for the manual and semi-automated methods (automated annotation did not include character emojis in the lexicon). This was done as tweets containing such characters clearly belonged to either the positive or negative categories. All tweets containing positive or negative character emojis were assigned as 'positive' or 'negative' respectively by the manual method indicating a complete success rate of allocating these tweets into the correct emotion category. Compared to this, the ability of the semi-automated method was less successful. 54.4% of tweets containing positive character emojis were misallocated by the semi-automated method as either 'neutral' or 'negative'; while 75% of the tweets containing negative character emojis were misallocated as 'neutral' or 'positive'.

3.4. Assignment of Tweets into Discrete Emotion Categories

Using the manual and semi-automated methods of annotation it was possible to assign tweets into a number of emotion categories. A comparison of the number of tweets assigned into each of these categories again highlights substantial variation between the methods (Figure 5). Both methods showed variation in the number of tweets they identified as belonging to each emotion category. Substantially higher numbers of tweets were annotated as anger/disgust, fear and beauty by the manual method compared to the semi-automated method.

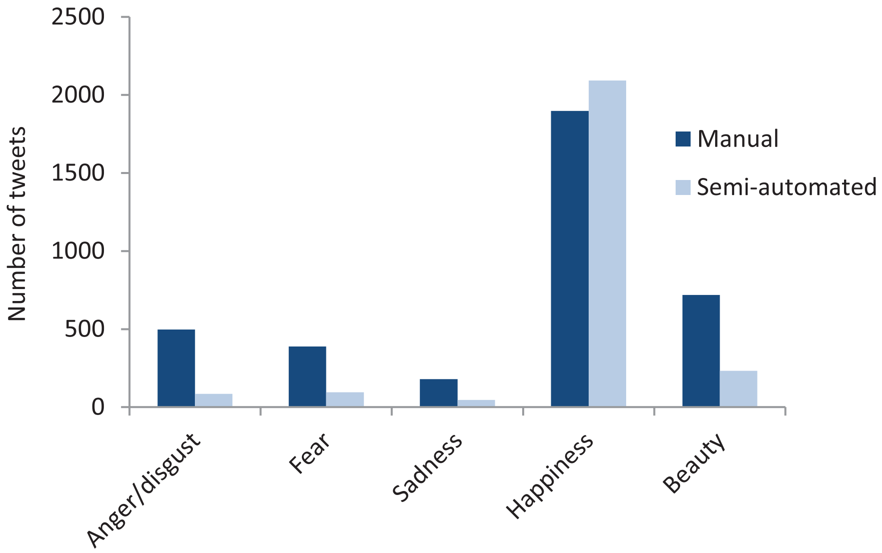


Figure 5. The number of tweets assigned by the manual and semi-automated methods into discrete emotion categories.

Percentage agreement between the two methods was found to be 44.5% when undertaken on all tweets. However, when tweets which were allocated as 'neutral' by both methods were removed, this figure falls substantially to 3.91%. This indicates that the methods show higher levels of variance when allocating an emotion to a tweet as opposed to just identifying the presence of an emotion, and that the presence of neutrality in a dataset can affect how the results of agreement between the assignment of tweets can first appear. A Cohen's Kappa Index of 0.0157 further emphasises the low level of agreement in allocation of tweets to discrete emotions.

4. Discussion

4.1. Comparison of the Outputs of Manual, Automated and Semi-Automated Analysis

The results presented show that detecting sentiments from tweets is a highly complex task, and importantly, that the method of analysis employed determines the categorisation of positivity, neutrality or negativity, despite using the same corpus of tweets. Moreover, the comparison of the manual and semi-automatic methods illustrated considerable variability in Ekman's specific emotion classes.

All three methods were found to assign variable yet similar numbers of tweets into the positive, neutral and negative categories, with the majority of tweets being annotated as neutral, followed by smaller numbers of positive and negative tweets respectively. Despite this analysis suggesting similarities between the three methods, assessment of inter-method reliability found percentile agreement between the assignment of tweets into the three categories by the methods to be only 444.5%.

The adjustment of thresholds used to assign automated tweet scores into positive, neutral and negative categories improved the similarity in the number of tweets assigned to each category between the manual and fully automated methods; however, it did not improve the percentage agreement between the two methods.

Manual annotation has previously been cited as providing the most reliable method of sentiment analysis, given that human annotators have the best chance of identifying the emotion present in a tweet (Saif, Fernez, He, & Alani, 2013). However, a dataset resulting from manual annotation is not unambiguous given that labelling tweets with an emotion remains a subjective task (Resch et al., 2016). Different human annotators may interpret the same text differently for many reasons—for example, sarcasm, slang or ambiguous use of emojis. This issue is also relevant for the semi-supervised learning method used here, given that the 'gold standard' tweet dataset used to train the algorithm relied on initial manual annotation of 1,000 tweets. To ensure that annotation was reliable between human annotators, a metric of comparison was derived suggesting agreement between them to be sufficient to assume inter-annotator reliability (Landis & Koch, 1977). Kappa Indexes enable the assessment of inter-annotator reliability between manual annotators and allow the variation in annotation by different annotators to be quantified.

Setting aside inherent subjectivity, the most significant limitation of manual sentiment analysis of tweets is the researcher time needed to examine each tweet. Given that Twitter generates large volumes of tweets in very short time periods, manual annotation is simply not viable. For this reason, automated and semi-automated methods are often employed.

Automated methods of sentiment analysis offer a quick and easy means of annotating large tweet datasets. Methodologically, however, there remains no robust way to derive discrete emotions from numeric scores, thus the granularity of the automated method demonstrated herein is limited to assessment of positivity rather than identifying specific emotions from tweet text. In this study, a large lexicon of words was used to enhance the reliability in the scores generated for each tweet. Despite this, the limitations seem to outweigh the benefits. Low inter-method reliability was prevalent and there was a particularly low percentage agreement between annotations of positive and negative suggesting that this method is unlikely to reliably identify the correct polarity of sentiment in tweet text. Additionally, while the large lexicon used provides robustness for scoring words, it does not include emojis which are increasingly common ways to express sentiment in short social media posts (Pavalanathan & Eisenstein, 2015). Previous research has shown that emojis can be successfully used to inform automated analysis of tweets (Go et al., 2009). Indeed, the creation of an emoji lexicon in which each is given a score would be of significant use to future research and enable the combined use of words and emojis in the annotation of sentiment from tweet text. Such an undertaken would need to overcome the challenge of interpreting emojis in their different representational forms:

Unicode (e.g. “U+1F642”), Kaomojis (e.g. “(◕‿◕)”), a sequence of ASCII characters (e.g. “:-)”) or a specific code used by Twitter (e.g., “<ed><a0><bd><ed><b2><af>” or “<ed><U+00A0><U+00BC><ed><U+00BC><U+009E>”).

An issue of spatial variation in language use was also identified associated with the automated method of annotation. Despite the large lexicon used, it cannot account for regional/local dialect. Given the location for this study was Birmingham, where some language used by local populations is not used elsewhere, these words will not have been included and scored and a proportion of sentiment in the tweets, albeit small, will not have been captured by this method. Provided that manual annotators are native to the language and region from which the tweets have been captured, this should not be an insurmountable issue.

The semi-automated method generated similar numbers of neutral, positive and negative tweets as the other two methods. However, Kappa Indices indicate that the placement of individual tweets into each of these categories showed low levels of agreement. Differences were also identified in how semi-automated annotation assigned tweets to discrete emotion categories, when compared to manual annotation. The notion of beauty is not a basic emotion as defined in emotion psychology; indeed, it is usually subsumed under happiness. This makes it difficult for the algorithm to identify beauty in written text because it is often expressed in comparatively subtle terms.

For the experiment presented in this article, it was possible to identify a limitation in the semi-supervised method, in that the full range of emojis in the dataset could not be captured by the algorithm. The method is designed for character-wise emojis (e.g. “:-)”), however unicode emojis are widely used alongside character-wise emojis in tweet texts. In fact, the semi-supervised learning method was not able to interpret unicode emojis, increasing the likelihood that essential elements of tweets were missed by this method, diluting the precision of assigning emotions and polarities.

The quality control measure, which used character emojis to assess the allocation of tweets into the correct category, highlighted that the semi-automated method was often unable to recognise emotion, despite these being included in the assessment of linguistic similarity undertaken during analysis.

The parameter choices of semi-automated approaches make such methods highly sensitive; the number of seeds used, the seed distribution, details of similarity computation, edge weight threshold and the emotion categories used strongly influence the results. A significant issue is that no formalised method exists to perform an *a priori* estimation for these parameters. In most cases, ‘optimal’ parameter settings can only be found through empirical experiments, which in turn means it cannot be stated with certainty how good any results are in relation to the best achievable results. Thus, the parameter choices require a substantial amount of expert knowledge and experience, particularly because random permutations cannot be performed due to the computational complexity of the algorithms. This opens up debate as to how a training dataset should be generated. In this article, 1,000 tweets were randomly chosen. It may be more appropriate to actively identify tweets which cover all the discrete emotion categories so the algorithm can learn most effectively.

Finally, in this article, for all the methods of emotion annotation used, it was assumed that one tweet contains a maximum of one emotion. However, in reality tweets can be inherently more complex and contain a variety of emotions over a short space of characters. This is a finding that future methods looking to classify the emotion in tweet text will need to consider and overcome to provide the most accurate interpretation of the emotional information that tweets contain.

4.2. Implications of These Findings for Urban Planners

The availability of emotional data to urban planners has significant utility in the creation, management and justification of urban green spaces which promote positive emotional experiences and minimise features which may elicit negative emotional responses (Roberts et al., in press). The provision of such emotional data through social networks, such as Twitter, provides the opportunity for planners to gain access to this information in inexpensive, time efficient and replicable ways. However, in order to be used meaningful, methodologies are required which can accurately annotate any emotion present in a tweet relating to an urban green space.

This article has identified that challenges remain to this end. Indeed, none of the three methods presented herein are appropriate in their current form to provide sentiment analysis of tweet text for urban planners.

Whilst manual analysis can be used to accurately identify any emotion present, the amount of time taken to undertake this method on a large corpus of tweets makes it unsuitable in the context of urban planning where resources and individuals are often limited.

Similarly, the current inability of automated and semi-automated methods to accurately identify emotion, make them dubious approaches to employ where the identification of such emotion and their causes could have significant implications for the management and creation of green spaces.

However, the authors tentatively suggest that pursuing a semi-automated method, like the one presented herein is the most appropriate way forward. The development of a method through which the accuracy of manual annotation can be achieved, in much shorter time is doubtless of interest to urban planners. This is of particular relevance because manual annotation of tweets is a time-consuming and expensive method. This article suggests that the development of a gold standard training data set should be a priority, enabling algorithms to learn the variety and complexity with which emotions can be conveyed in tweets.

Without a doubt, Twitter data presents a useful and abundant source of easily accessible emotion information which is generated by users as they experience specific urban green spaces. Such a source of data presents vast opportunities for urban planners; however there remains a need for increased innovation and development in the methodologies which would enable this data source to be engaged with most effectively.

5. Conclusion

This paper has presented a comparison of three approaches to sentiment analysis undertaken to collate the sentiment and emotion present in tweet text. Despite their utility, significant differences exist in the outcomes of three methods of sentiment analysis on the same corpus of tweets. The discrepancies in how tweet text is analysed by different methods is thus a critical consideration for future research.

It was possible to identify differences in positivity annotation between all three methods in terms of the numbers of tweets assigned to each category as well as inter-method reliability in assignment. Using the manual and semi-automated methods, discrete emotions can be annotated, but again significant differences were identified in this process, particularly for beauty and anger/disgust tweets.

Overall, whilst this article is positive about the role of Twitter in providing a useful and substantial data source for urban planners on which to undertake sentiment analysis, it suggests caution is needed in interpreting the outputs of sentiment analysis and an understanding of the process can help place the results in an appropriate context. A critical discussion of the limitations identified through the undertaking of all three methods in this research has been presented. In doing so, it adds to the debate surrounding annotation of sentiment and emotion from tweets and identifies methodological constraints which should be taken into account in future work. Given the utility of the sentiment information captured by tweets relating to urban green space for planners and decision makers, it is of important that an efficient and reliable method is established through which these can be identified and annotated. Despite its reliability, manual annotation is unfeasible for large volumes of data. However, automated and semi-automated methods are hampered by a number of limitations associated with each, and this work shows that methodological progression is necessary before either can be used robustly to annotate sentiments from large tweet datasets.

The findings presented here suggest that automated methods of sentiment analysis are not able to accurately identify the emotion present in tweet text and that manual analysis, whilst accurate, is impractical for use on large tweet corpi given the time taken to undertake such analysis. As a result, this research suggests that future attempts to develop methods of sentiment analysis should focus on semi-automated methods, with particular focus given to how the gold standard dataset is selected. Successful algorithms should aim to include Unicode as well as character emojis in order to best capture the emotion represented by these in tweets.

Acknowledgments

We would like to express our gratitude to the Austrian Science Fund (FWF) for supporting the project “Urban Emotions”, reference number I-3022. We would also like to thank Dr. Wendy Guan from Harvard University’s Center for Geographic Analysis for her support through providing us with the Twitter data for our study. We would also like to thank the Engineering and Physical Sciences Research Council (EPSRC), reference number EP/L504907/1.

Conflict of Interests

The authors declare no conflict of interests.

References

Barbosa, L., & Feng, J. (2010). Robust sentiment detection on twitter from biased and noisy data. In *Proceedings of the 23rd International Conference on Computational Linguistics: Posters* (pp. 36–44). Stroudsburg, PA: Association for Computational Linguistics.

Bertrand, K., Bialik, M., Virdee, K., Gros, A., & Bar-Yam, Y. (2013). Sentiment in New York City: A high resolution spatial and temporal view. *arXiv preprint arXiv:1308.5010*

Birmingham City Council. (2016). Parks and nature conservation. Retrieved from http://www.birmingham.gov.uk/parks

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, *2*(1), 1–8.

Bruns, A., & Burgess, J. (2011). The use of Twitter hashtags in the formation of ad hoc publics. In *Proceedings of the 6th European Consortium for Political Research (ECPR) General Conference 2011*.

Cheong, M., & Lee, V. (2011). A microblogging-based approach to terrorism informatics: Exploration and chronicling civilian sentiment and response to terrorism events via Twitter. *Information Systems Frontiers*, *13*(1), 45–59.

Chung, S., & Liu, S. (2011). *Predicting stock market fluctuations from twitter*. Retrieved from https://www.stat.berkeley.edu/ aldous/157/Old_Projects/Sang_Chung_Sandy_Liu.pdf

Cohen, D., Sehgal, A., Williamson, S., Marsh, T., Golinelli, D., & McKenzie, T. (2009). New recreational facilities for the young and old: Policy and programming implications. *Journal of Public Health Policy*, *30*(1), S248–S263.

Costanza, R., d'Arge, R., Groot, R., Farber, S., Grasso, M., Hannon, B., . . . Belt, M. (1997). The value of the world's ecosystem services and natural capital. *Nature*, *387*, 253–260.

Daily, G. (Ed.). (1997). *Nature's services: Societal dependence on natural ecosystems*. Washington, DC: Island Press.

Daniel, T., Muhar, A., Arnberger, A., Aznar, O., Boyd, J., Chan, K., . . . von der Dunk, A. (2012). Contributions of cultural services to the services agenda. *PNAS*, *109*(23), 8812–8819.

Dodds, P., Harris, K., Kloumann, I., Bliss, C., & Danforth, C. (2011). Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PLOS ONE*, *6*(12), e26752.

Ehrlich, P., & Ehrlich, A. (1981). *Extinction: The causes and consequences of the of the disappearance of species*. New York: Random House.

Ekman, P. (1999). Basic emotions. In T. Dalgleish & M. Power (Eds.), *Handbook of cognition and emotion* (pp. 45–60). New Jersey: Wiley and Sons Ltd.

Ekman, P., & Friesen, W. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, *17*(2), 124.

Fuller, R., & Gaston, K. (2009). The scaling of green space coverage in European cities. *Biology Letters*, *5*(3), 352–355.

Go, A., Bhayani, R., & Huang, L. (2009). Twitter sentiment classification using distant supervision. *CS224N Project Report, Stanford*, *1*(12).

Grimm, N., Grove, J., Pickett, S., & Redman, C. (2000). Integrated approaches to longterm studies of urban ecological systems. *Bioscience*, *50*, 571–584.

Hollander, J., Graves, E., Renski, H., Foster-Karim, C., Wiley, A., & Das, D. (2016). *Urban social listening: Potential and pitfalls for using microblogging data in studying cities*. Basingstoke: Palgrave Macmillan.

Jansen, B., Zhang, M., Sobel, K., & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *Journal of The American Society for Information Science and Technology*, *60*(11), 2169–2188.

Kellert, S., & Wilson, E. (1995). *The biophilia hypothesis*. Washington, DC: Island Press.

Landis, J., & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *1977*, 159–174.

Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, *5*(1), 1–167.

Mandel, B., Culotta, A., Boulahanis, J., Stark, D., Lewis, B., & Rodrigue, J. (2012). A demographic analysis of online sentiment during hurricane Irene. In *Proceedings of the Second Workshop on Language in Social Media* (pp. 27–36). Stroudsburg, PA: Association for Computational Linguistics.

Martinez, V., & González, V. (2013). Sentiment characterization of an urban environment via Twitter. In *Ubiquitous computing and ambient intelligence. Context-awareness and context-driven interaction* (pp. 394–397). Berlin: Springer Verlag.

McHugh, M. (2012). Interrater reliability: The kappa statistic. *Biochem Med*, *22*(3), 276–282.

MEA. (2005). *Ecosystems and human well-being* (Vol. 5). Washington, DC: Island Press.

Milcu, A., Hanspach, J., Abson, D., & Fischer, J. (2013). Cultural ecosystem services: A literature review and prospects for future research. *Ecology and Society*, *18*(3), 44.

Mitchell, L., Frank, M., Harris, K., Dodds, P., & Danforth, C. (2013). The geography of happiness: Connecting twitter sentiment and expression, demographics, and objective characteristics of place. *PLOS ONE*, *8*(5), e64417.

Office for National Statistics. (2014). *Mid-2014 mid-year population estimates, BDB2015/04 Birmingham demographic briefing*. London: ONS.

Pak, A., & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. *LREC*, *10*.

Pavalanathan, U., & Eisenstein, J. (2015). Emoticons vs. emojis on Twitter: A causal inference approach. *arXiv preprint arXiv:1510.08480*

Quercia, D., Ellis, J., Capra, L., & Crowcroft, J. (2012). Tracking gross community happiness from tweets, In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (pp. 965–968). New York: ACM.

Resch, B., Summa, A., Zeile, P., & Strube, M. (2016). Citizen-centric urban planning through extracting emotion information from Twitter in an interdisciplinary space-time linguistic algorithm. *Urban Planning*, *1*(2), 114–127.

Roberts, K., Roach, M., Johnson, J., Guthrie, J., & Harabagiu, S. (2012). EmpaTweet: Annotating and detecting emotions on Twitter. *LREC*, 3806–3813.

Roberts, H., Sadler, J., & Chapman, L. (in press). The value of Twitter data for determining the emotional responses of people to urban green spaces: A case study and critical evaluation. *Urban Studies*.

Rosenthal, S., Ritter, A., Nakov, P., & Stoyanov, V. (2014). Semeval-2014 task 9: Sentiment analysis in twitter. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)* (pp. 73–80).

Saif, H., Fernez, M., He, Y., & Alani, H. (2013). Evaluation datasets for Twitter sentiment analysis: A survey and a new dataset, the STS-Gold. In *1st International Workshop on Emotion and Sentiment in Social and Expressive Media: Approaches and Perspectives from AI (ESSEM 2013)*, Turin, Italy.

Shalunts, G., Backfried, G., & Prinz, P. (2014). Sentiment analysis of German social media data for natural disasters. In *Proceedings of the 11th International Conference on Information Systems for Crisi Response and Management (ISCRAM)*.

Shanahan, D., Fuller, R., Bush, R., Lin, B., & Gaston, K. (2015). The health benefits of urban nature: How much do we need? *Bioscience*, biv032.

Shanahan, D., Lin, B., Gaston, K., Bush, R., & Fuller, R. (2014). What is the role of trees and remnant vegetation in attracting people to urban parks? *Landscape Ecology*, *30*(1), 153–165.

Talukdar, P., & Pereira, F. (2010). Experiments in graph-based semi-supervised learning methods for class-instance acquisition. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics* (pp. 1473-1481). Uppsala, Sweden, Association for Computational Linguistics.

Thelwall, M. (2014). Sentiment analysis and time series with Twitter. In K. Weller, A. Bruns, J. Burgess, M. Mahrt, & C. Puschmann (Eds.), *Twitter and society* (pp. 83–95). New York: Peter Lang.

Tumasjan, A., Sprenger, T., Sandner, P., & Welpe, I. (2010). Predicting elections with twitter: What 140 characters reveal about political sentiment, *ICWSM*, *10*(1), 178–185.

Turney, P. (2002). Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics* (pp. 417–424). Stroudsburg, PA: Association for Computational Linguistics.

Ulrich, R. (1983). Aesthetic and affective response to natural environment. In *Behavior and the natural environment*. Boston, MA: Springer.

UN Habitat. (2016). *World city report 2016. Urbanisation and development: Emerging futures*. Nairobi: UN Habitat.

Wang, H., Can, D., Kazemzadeh, A., Bar, F., & Narayanan, S. (2012). A system for real-time twitter sentiment analysis of 2012 US presidential election cycle. In *Proceedings of the ACL 2012 System Demonstrations* (pp. 115–120). Stroudsburg, PA: Association for Computational Linguistics.

Warriner, A., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal and dominance for 13,915 English lemmas. *Behavioural Research*, *45*, 1191–1207.

Wilson, E. (1984). *Biophilia*. Boston: MA: Harvard University Press.

World Health Organisation. (2017). *Urban green spaces: A brief for action*. Denmark: WHO Regional Office for Europe.

Zhang, L., Riddhiman, G., Dekhil, M., Hsu, M., & Liu, B. (2011). *Combining lexicon-based and learning-based methods for Twitter sentiment analysis* (HPL-2011-89). Palo Alto, CA: HP Laboratories.

About the Authors

Helen Roberts is a Doctoral Researcher at the University of Birmingham using social media data to investigate a variety of urban interactions between people and green spaces, particularly parks and corridors. A socio-ecological perspective frames this work and examines key how certain aspects of ecosystem service provision are dependent on human use of green spaces.

Bernd Resch is an Assistant Professor at University of Salzburg's Department of Geoinformatics—Z_GIS and a Visiting Fellow at Harvard University (USA). His research interests revolve around fusing data from human and technical sensors, including the analysis of social media. Amongst a variety of other functions, Bernd Resch is Editorial Board Member of the *International Journal of Health Geographics* and the *International Journal of Geo-Information*, Associated Faculty Member of the doctoral college "GIScience", and Executive Board member of Spatial Services GmbH.

Jon Sadler is a Biogeographer and Ecologist whose research focuses on species population and assemblage dynamics in animals (sometimes plants). His work is highly interdisciplinary, bisecting biogeography, ecology, urban design, riparian management and island Biogeography. It uses approaches that combine detailed field studies, field and laboratory experimentation, sometimes with social science to examine the links between environmental variability and species (including humans) responses. His research has significant blue skies and applied implications for understanding and responding to the impacts of climate and environmental change variability on urban and island ecosystems, hydrological systems, riparian/riverine ecology, the management/conservation of freshwaters. He is a fan of numbers and coding (especially using open source software such as R).

Lee Chapman is a Professor of Climate Resilience at the University of Birmingham focussed on researching the impact of weather and climate on the built environment. His research has a particular emphasis on critical infrastructure systems in urban areas and includes significant knowledge transfer to the meteorological market place.

Andreas Petutschnig received his BEng in Cartography and Geomedia-technology at the University of Applied Sciences in Munich, Germany and his MSc in Applied Geoinformatics at the University of Salzburg, Austria. His research interests include the analysis and visualisation of spatiotemporal data flows, spatial statistics, scaling issues of spatial data clusters, and reproducible research. His PhD work has a focus on spatiotemporal point pattern analysis. Specifically, this includes the detection of refugee camps in crisis regions based on traces of social media data.

Stefan Zimmer studied Geoinformatics (BSc) in Germany and currently finalises the Masters progamme in Geoinformatics. In the research group, he is involved in building and improving the social media crawler infrastructure. Furthermore, Stefan optimizes the performance of the emotion classification algorithm for social media posts by exploiting massive parallelism on graphic processing units with Set Similarity Joins.

Urban Planning (ISSN: 2183–7635)
2018, Volume 3, Issue 1, Pages 34–46
DOI: 10.17645/up.v3i1.1261

Article

All Work and No Play? Facilitating Serious Games and Gamified Applications in Participatory Urban Planning and Governance

Cristina Ampatzidou [1,]*, Katharina Gugerell [1], Teodora Constantinescu [2], Oswald Devisch [2], Martina Jauschneg [3] and Martin Berger [3]

[1] Department of Spatial Planning & Environment, University of Groningen, 9747AD Groningen, The Netherlands; E-Mails: c.ampatzidou@rug.nl (C.A.), k.gugerell@rug.nl (K.G.)
[2] Faculty of Architecture & Arts, Hasselt University, 3590 Diepenbeek, Belgium; E-Mail: teodora.constantinescu@uhasselt.be (T.C.), oswald.devisch@uhasselt.be (O.D.)
[3] Green City Lab Vienna, 1010 Vienna, Austria; E-Mails: office@jauschneg.at (M.J.), martin.kp.berger@tuwien.ac.at (M.B.)

* Corresponding author

Submitted: 14 November 2017 | Accepted: 22 January 2018 | Published: 29 March 2018

Abstract
As games and gamified applications gain prominence in the academic debate on participatory practices, it is worth examining whether the application of such tools in the daily planning practice could be beneficial. This study identifies a research–practice gap in the current state of participatory urban planning practices in three European cities. Planners and policymakers acknowledge the benefits of employing such tools to illustrate complex urban issues, evoke social learning, and make participation more accessible. However, a series of impediments relating to planners' inexperience with participatory methods, resource constraints, and sceptical adult audiences, limits the broader application of games and gamified applications within participatory urban planning practices. Games and gamified applications could become more widely employed within participatory planning processes when process facilitators become better educated and better able to judge the situations in which such tools could be implemented as part of the planning process, and if such applications are simple and useful, and if their development process is based on co-creation with the participating publics.

Keywords
citizen engagement; games; gamification; participatory planning; serious games; urban governance; urban planning

Issue
This article is part of the issue "Crowdsourced Data and Social Media for Participatory Urban Planning", edited by Bernd Resch (University of Salzburg, Austria), Peter Zeile (Karlsruhe Institute of Technology, Germany) and Ourania Kounadi (University of Salzburg, Austria).

1. Introduction

Games and gamified applications are often described as being a magic bullet in current governance debates, with their aim to attract citizens to engage with city matters and planning questions, to participate in decision-making, and to improve the overall process of public participation. Public engagement is dominated by conceptual and practical difficulties, it is still framed in the dominant rhetoric of mainly involving the citizens who are 'affected' by the plans, and it takes place within timeframes set by the respective planning procedures and contracting organisation (Horelli, 2002). Thus, an increasing number of people perceive participation as pointless and rarely able to resolve conflicts or influence decision-making (Innes & Booher, 2010). Instead of attending another community meeting people would rather spend their 'leisure-time' on activities they appreciate and truly enjoy (Lerner, 2014). Other authors emphasise that citizens still engage but the ways of communication have changed drastically, complementing and partly even replacing community meetings and co-located participatory action with digital tools and social media (Ekman & Amnå, 2012; Gordon & Mihailidis, 2016; Hay, 2007; Kleinhans, van Ham, & Evans-Cowley, 2015; Macafee & De Simone, 2012; Marichal, 2013; Skocpol, 1997; Tufekci & Wilson, 2012). In different planning and design-related disciplines, digital tools for online participation, such as e-democracy portals, online consultations, e-voting, crowdsourcing, blogging, social network platforms, mobile apps, community GIS, and online deliberation, have gained increased attention as instruments to involve citizens and actor groups who either are too busy or not interested to participate in co-located meetings (e.g., Ahmed, Mehdi, Moreton, & Elmaghraby, 2015; Belluci et al., 2015; Kelley & Johnston, 2011; Prandi, Roccetti, Salomoni, Nisi, & Jardim Nunes, 2017).

Such tools have raised expectations of the potential to overcome barriers to public participation such as the lack of long-term engagement, inclusion, and empowerment of underrepresented actor groups, as well as more broadly to make the process of public participation more pleasant and enjoyable. The vivid academic debate on the development and benefits of novel formats and tools, especially on games and gamified environments and tools (e.g., Devisch, Poplin, & Sofronie, 2016; Medema, Furber, Adamowski, Zhou, & Mayer, 2016; Poplin, 2014; Tolmie, Chamberlain, & Benford, 2014), strongly focuses on the 'supply side' of the issue. The term 'supply' addresses the conceptual framing, design and development, and experimental testing of serious games and tools in public participation and urban governance, typically within the context of a research project or a living lab. Less attention is paid to the 'demand side': the investigation of the actual practises, experiences, expectations, and barriers to implementing and facilitating such tools in the daily, regular work practice. In this article we target this gap, by investigating the experiences of planning professionals with novel formats, namely games and gamified applications in their daily practice, addressing the following research questions: (i) which formats actually form part of the daily practice in participatory urban planning processes?; (ii) which benefits and advantages do facilitators identify in working with such formats?; and (iii) what are barriers which impede the willingness and ability to work with and facilitate such formats?

The article is organised as follows: in Section 2 we introduce the academic debate around games and gamified applications in addressing long-standing challenges relating to civic participation in urban planning and outline how games and gamified applications have been used in urban planning and governance so far. Section 3 describes the methodological approach and the three case study cities. In Section 4 we present and discuss our findings regarding the current state of participation in the three case study cities as well as the perceived benefits and obstacles in employing games and gamified applications within participatory planning settings. Finally, in Section 5 we return to our research objectives to conclude that games and gamified applications could claim a larger share of the tools employed within participatory planning processes, when their development process is based on co-creation with the participating publics, when they are simple and developed with careful use of the available resources, and when process facilitators are better educated and better able to judge the situations in which such tools could be implemented as part of the planning process.

2. Serious Games and Gamified Applications in Participatory Planning Practice

Civic engagement and citizen participation can be broadly defined as the sum of political and social practices, by which individuals engage with and influence public affairs, beyond their direct private environment (Gordon, Balwin-Philippi, & Balestra, 2013; Parés & March, 2013; Raphael, Bachen, Lynn, Balwin-Philippi, & McKee, 2010). Engagement and participation has become an inherent part of urban planning and governance, and is facilitated by different tools and methods, well beyond its traditional expressions of voting and attending town hall meetings (Gordon & Mihailidis, 2016). Participatory methods are used to address a variety of aspects in urban planning and architectural design, including design issues, stakeholder negotiations and deliberation, and enabling self-organisation (Glick, 2012; Grahan & Marvin, 2001; Krasny, 2013; Uitermark & Duyvendak, 2008). Experimenting with novel tools and technologies, such as mobile apps, social media, games and gamified environments are efforts to both, diversify the media used for civic engagement, support the creation of different results, and at least partially address persistent common underlying problems (Rowe & Frewer, 2000; Shipley & Utz, 2012), such as the often-downplayed undercurrent of social conflict and power struggles (Fainstein, 2000; Sandercock, 1994), the inequality of bargaining power among various stakeholders (Lane, 2005) or deal-brokering behind closed doors (Innes & Booher, 2004), the overrepresentation of the so-called usual suspects and extreme viewpoints (Fiorina, 1999); the difficulties of including socioeconomically disadvantaged groups, the lack of expertise and motivation among citizens, high drop-out rates, as well as the lack of trust in the government's ability to make good use of the participatory processes (Brown & Chin, 2013; Irvin & Stansbury, 2004; Tonkens, 2014). A broad range of digital media and tools enter the field of civic participation because of their ability to incorporate larger volumes of data and information of different types (visual, textual, sound, etc.) and to present them in user-friendly formats to raise awareness and engage a broader audience (Gramberger, 2001; Kleinhans et al., 2015). Providing information and participation opportunities for distributed and remote citizens has also entered governmental offices and public bodies, often resulting in the establishment of 'innovation offices' responsible for an increasing number of digital online services and newly developed or redesigned, more easily accessible engagement tools (Conroy & Evans-Cowley, 2005; Gordon & Mihailidis, 2016). Even though the 'supply' of new formats demonstrates an extensive variety of new tools, the experimentation with and adoption of novel participatory formats by the 'demand side' is not straightforward. Many planners address the lack of sufficient education and training in participatory methods (Ekman & Amnå, 2012; Handley & Howell-Moroney, 2010; Innes & Booher, 2004). Others report on the limitations posed by existing regulatory frameworks which enforce the use of specific methods and fail to follow the pace of technological development of innovative engagement tools (Houghton, Miller, & Foth, 2014). Within public administration, lack of time, knowledge, and desire are also debated as being important reasons for non-participation (Yang & Callahan, 2007).

In urban planning, the use of games in particular has a profound history since the 1960s (Abt, 1969; Duke, 1975), and has remained a popular tool for spatial modelling and simulation, and public participation (Devisch et al., 2016; Mayer, 2009; Poplin, 2012). Even though a universally shared definition of what constitutes a 'game' is lacking, there is agreement that games are a form of structured play (Salen & Zimmerman, 2004). That means that games include sets of rules that enable and restrain the players' pursuit of a predetermined goal. 'Serious games' is a more recent field of game studies focusing on games that also include educational goals, instead of exclusively being for amusement (Abt, 1969). Early applications of serious games in urban planning focused mainly on ways to overcome challenges on the level of understanding and modelling urban dynamics, addressing topics such as land use, transportation, ecology, and management of natural resources. One of Abt's first urban games was 'Corridor' (Abt, 1969), a computer-assisted simulation game, to explore the technological, economic and political constraints on the development of an alternative transportation plan for the Northeast Corridor, between Boston and Washington D.C. In the 1960's and 70's Jay Forrester's (1969) work on urban dynamics inspired a series of urban simulation games, such as the games developed by Meadows and Randers for the Club of Rome, and even the popular city-building game SimCity (Mayer, 2009). 'Climate Hope City' (Blockworks, 2015) and 'Port of the Future' (Deltares, 2016) are contemporary simulation games, addressing challenges of resource management, urban power grid simulations, renewable energy and decision making. Even though simulation and modelling still play a pivotal role in urban planning and policy games, the potential of games to create environments for learning, negotiation, deliberation and collaboration among players is attracting increasingly more attention, which is also informed by the rising interest in gamification (Devisch et al., 2016; Gordon & Baldwin-Philippi, 2014; Poplin, 2012; Tan, 2014). Many recent games provide ample opportunities for analogue and digital social interaction among players. The DuBes Game (van Bueren, Mayer, Bots, & Seijdel, 2007), for example, is explicitly organised around two workshop sessions where players assume different stakeholder roles and negotiate an agenda for sustainable urban renewal. 'Age of Energy' (Clicks and Links, 2015) in an app-based game where players compete against their neighbours to save energy in real life. In such games, we ascertain a shift of focus from spatial understanding towards social aspects of playing in hybrid game-real-world settings.

Deterding et al. (2011) stress the importance of a strict distinction between gamification, and (serious) games. While games are considered to trigger the experiential and behavioural qualities of gameplay, gamified applications are notably centred around the use of specific game elements invoking gameful (ludic) qualities (Deterding, Dixon, Khaled, & Nacke, 2011). Gamification describes "the use of game design elements in non-game contexts" (Deterding, Dixon, Khaled, & Nacke, 2011). Gamification came to prominence in the early 2010's, mainly as enhanced advertising and marketing practices, where game elements such as points, badges and leaderboards were used to motivate audiences to engage with certain applications or brands (Huotari & Hamari, 2011; Lindqvist, Cranshaw, Wiese, Hong, & Zimmerman, 2011; Zichermann & Cunningham, 2011). Gamification has been contested by several researchers especially within game studies, as reducing games to simple point collection (Robertson, 2010), as a form of exploitation (Bogost, 2011, 2014), and as a face-saving mechanism preventing deeper engagement (DiSalvo & Morrison, 2011). Gamification is often applied in participatory urban planning by using game elements to enable citizens to debate or give feedback on specific plans and to propose ideas for small-scale projects. For example, in Participatory Chinatown (Gordon, 2010), citizens were able to virtually walk around Boston's Chinatown and comment on the proposed developments. In 'Neighborland' (Parham, Parham, & Chang, 2011) civic organizations can inform and engage citizens about their projects, run surveys, and ask people to comment and propose ideas.

The interest of urban scholars in serious games and gamified applications stems from games' specific ability to balance entertainment and learning (Abdul Jabbar & Felicia, 2015; Boyle, Connolly, Hainey, & Boyle, 2012; Whitton, 2011). The learning aspects of gaming have more recently been associated with a series of benefits for participation and civic engagement, such as raising awareness, increasing literacy on specific topics, developing (complex) problem solving skills, the ability to test difficult scenarios within a safe environment, and to establish networks and coalitions (e.g., Crookall, 2010; Erhel & Jamet, 2013; Gee, 2005; Granic, Lobel, & Engels, 2014; Luederitz et al., 2016; Shaffer, Squire, Halverson, & Gee, 2005). In this article, we investigate how experimentation with games and gamified applications takes place in participatory urban planning practice, the benefits as seen by planners and facilitators, as well as barriers which they are confronted with in their daily work.

3. Methodological Approach and Introduction to the Case Studies

The experiences and expectations of urban planners in using serious games and gamified applications to support participatory urban planning practices were studied in the cities of Groningen (NL), Vienna (AT) and Genk (BE). The case study selection is based on the research project that this work is embedded in. The three cities cover a broad spectrum of spatial and socio-cultural settings. Even though we expected to see diverse applications of participatory processes due to the different institutional, spatial, socio-economic and cultural conditions, and the broad variety of experiences and expectations resulting from the broad cultural and institutional contexts, we were able to combine the observations from the three cities due to the commonalities observed regarding the organisation and facilitation of participatory processes as well as the tools that were used during such processes (Table 1).

The article is based on an explorative case study research (Scholz & Tietje, 2002; Yin, 2013) following a two stages approach, combining an initial online explorative survey with guideline-based expert interviews. The survey was used to collect background information, to systematically map the participatory approaches, tools and methods currently in use within planning processes, as well as to identify which topics were addressed, which audiences were included, and the perceived impact of participation on the planning processes. The survey also covered the participants' general experience with games and gamified applications in the three cities and delivered the basis for the interview guidelines. Interviews were carried out in English, in Groningen and Genk, and in German in Vienna and all were based on the same guidelines. The expert interviews (Groningen: 8, Vienna: 7, Genk: 9), covered professionals within the fields of public administration (10 interviews), research

Table 1. Overview of the three case studies and the participatory methods used in Vienna, Genk and Groningen, based on the expert interviews and document analysis.

City	Vienna (AT)	Genk (BE)	Groningen (NL)
Current focus regarding planning & development	Urban planning, community development, mobility, energy transition, carbon footprint, local economy		
Process planning & design	District service, area renewal office, agenda 21 office	Neighbourhood management (Wijk management)	Process management, public servants
Implementation: methods facilitated	Large variety: focus on traditional, well-established methods, like focus groups or brainstorming techniques, partly extended by social media platforms		
	Workshops, brainstorming techniques, focus groups, public interventions	Brainstorming techniques, meetings/discussion rounds	Meetings, discussion rounds, information distribution
Resource restrictions on participant level	Time, knowledge, language barriers, educationally deprived groups & low-income groups, cultural restrictions (hard-to-reach-groups)	Time, knowledge, language barriers, cultural restrictions	Knowledge, know how/technical capacity, language barriers (present but not perceived)
Diversity	Underrepresentation of non-European groups, adults/working population (well represented) and elderly people (65+) tendency towards over-representation		
Digital tools	Participatory GIS, mobile apps, quizzes, online forums and feedback forms, tv, radio, project website	Photography, filming, quizzes, online forums, tv, radio, project website	Surveys (online), social media monitoring, photography, filming, online reaction forms, tv, radio, project website
Games and gamified applications	Board games, explorative board games for idea development, role playing, street games	Educational games, games in a business context, urban games, story-collecting birds	Gamified participatory GIS, city development game

(3 interviews) and facilitators of public participation processes, such as civic engagement offices or district renewal agencies (11 interviews). The expert interviews made an in depth exploration of the variety of participatory projects and engagement processes that the interviewees employed in their daily practices, the perceived value of using participatory processes, the diversity of methods and (digital) tools that were used, as well as the problems they encountered. Based on their previously declared familiarity with games and gamified applications, the interviews explored either their experiences of employing such tools and their (positive or negative) evaluation of the reasons for not engaging with alternative formats, as well as their expectations. The interviews were transcribed, coded and analysed using qualitative content analysis (Gläser & Laudel, 2010; Mayring, 2015).

Participatory processes and tools have been analysed using Horelli's (2002) methodological scheme of participatory planning, conceptualizing it as an evolutionary process that unfolds over time and consists of five generic steps, in which multiple tools can be used to facilitate communicative transactions among participants in specific environmental, organizational, economic, cultural and temporal contexts: (i) initiation of the project, (ii) planning and design, (iii) implementation, (iv) evaluation and research, and (v) maintenance (Horelli, 2002). These phases are interconnected by a continuous monitoring, providing feedback on the progress, quality and results of the process. Facilitating tools are structured in four categories based on their potential to enable communicative transactions: (i) diagnostic tools enable the determination of existing resources, mapping of the context and definition of the desired outcomes of the process; (ii) expressive tools enable participants to communicate their ideas and express themselves; (iii) organizational tools are those that underlie the organization of the process, including the creation of events, and (iv) political tools address common goal setting and power differences (Horelli, 2002).

4. Current State of Use of Games and Gamified Applications in Participatory Settings in Vienna, Groningen and Genk

The identified commonalities among the three case study cities notably surpass their differences in the scope of tools used in participatory settings (Tables 1 and 2). These commonalities allow the establishment of the 'demand side' regarding employment of games and gamified applications. Most participatory projects mentioned by the interviewees were linked to urban planning, infrastructure development, community development and the local economy, and were initiated and commissioned by a governmental organisation. The projects also varied in scale, ranging from street level to neighbourhood and city-wide, as well as infrastructure, urban policy, and urban design. For the most part, these are projects that are considered to be highly relevant to citizens' daily life, such as community building processes, projects directly linked to the (spatial) quality of the neighbourhood and quality of life, and big infrastructure projects that are expected to affect a large population over an extended period of time. Municipalities, groups of organisations and activist initiatives play an important role in launching topics, raising awareness, and initializing participatory projects. However, civic participation is often outsourced to intermediary organisations and planning agencies. Thus, the demand for new ways of engagement stems not only from the decision-making bodies, but also from these intermediaries and other initiating stakeholders.

A broad variety of tools and methods are already in use across all four categories of Horelli's (2002) framework (Table 2). A great deal of emphasis is placed on the initiation, and the planning and design phase, where the large majority of tools are concentrated. These are the most intensive phases of the participation process because they will enable the project to proceed smoothly. Comparatively little attention is paid to the evaluation phase, with a few instances of feedback being sought following the success of the projects, with fewer tools being used during the implementation and maintenance phases. These phases are often considered 'technical', in the sense of their being able to be carried out in a straightforward way by expert professionals, and thus are thought to not require broader public engagement. A large number of tools are classified as organizational: these are mostly tools that provide project information, information about the development process, and the state of works to the public. Exhibitions, guided tours, and leaflets tools are the only tools which have a significant presence during the implementation phase, these being tools which provide information but collect no feedback. The political category has the least number of tools available to it; there being very few tools used that address common goals and power differences of participating actors which points to the lesser importance given to questioning the predefined conceptual structures of the planning processes. Finally, regarding the nature of the tools, even though digital tools are consistently present throughout the process and across all categories, non-digital tools continue to dominate the daily participatory planning practice.

The facilitators and planners generally choose the tools and methods they feel the most comfortable with, resulting in the prevalence of non-digital methods and tools: "real games in the narrower sense are not used. Well, we have—we use more traditional methods, such as moderations and surveys and such" (VIE-JG). Even though the more 'traditional' formats implied in this quote remain the majority, experimentation with new media and digital tools also exists. The use of these tools happens either very early in the process with the aim to initiate and support an active and positive start of the participatory process (e.g., activation of participants, getting acquainted, capacity building on the planning pro-

Table 2. Overview of tools used in participatory processes in the three case study cities, following Horelli (2002).

TOOLS		PLANNING PROCESS				
		Initiation	**Planning & Design**	**Implementation**	**Evaluation & Research**	**Maintenance**
Diagnostic	Non-digital	Surveys (offline), interviews			Surveys (offline)	
	Digital	Surveys (online), social media monitoring, photography, filming, participatory GIS			Surveys (online), social media monitoring	
Expressive	Non-digital	Interviews, focus groups, consultation meetings, workshops, activation games, quizzes	Architectural models, interviews, focus groups, consultation meetings, games, workshops, brainstorming		Interviews	
	Digital	Mobile apps, games, quizzes	Mobile apps, online forums and feedback forms, games		Mobile apps, online forums and reaction forms	Online forums and feedback forms
Organizational	Non-digital	Local press, policy documents and reports, brochures and flyers, press conferences, letters	Guided tours, exhibitions, info points, on-site info panels, brochures and flyers, letters	Guided tours, exhibitions, info points, on-site info panels, brochures and flyers	Policy documents and reports, letters	
	Digital	Tv, radio, project website,	Project website	Project website	Project website	Project website
Political	Non-digital	Fund-raising	Participatory budgeting	Co-financing	Citizen panels	

cesses), or later in the process to produce content in the planning and design phase (e.g., developing a proposal for a park, strategy development for a harbour). As for the production of content, the focus is on using a variety of expressive tools for the development of planning proposals (e.g., neighbourhood parks and squares in Vienna), urban strategic plans (e.g., port redevelopment in Rotterdam), and for considering perspectives and wishes from various stakeholders and actors:

> We used it in a part of the former harbour, not so much for urban planning, more to get an urban strategy and to make a deal with all participants because there were private owners, the central Dutch government, the city, all kinds of parties who had some role in this whole area. (GRO-ES)

With regards to the process, we see that especially in early stages of participatory processes gamified applications are facilitated, targeting activation, allowing participants to become acquainted with one another, or developing the knowledge required to enter the participation process: "Rather 'activation-games' and also quizzes for knowledge creation about the carbon footprint or mobility" (VIE-SH). New media and digital tools are used regularly to motivate and inform participants, but also at later stages they can lower threshold for participation, acting as icebreakers and facilitating social interaction within the group of participants: "it's an extra way for people to get in, I think….It's a way of connecting with others" (GRO-AH) or "But you can see all kinds of groups processes going on….It's not only about the plan or the development but also about the interaction in the group" (GRO-ES).

4.1. Experienced Benefits of Games and Gamified Applications in Participatory Processes

We illustrated that so-called 'traditional' methods for civic engagement in participatory urban planning coexist with experimentation using novel media, tools, and games. Three main perspectives emerged from the case studies showing the perceived benefits of using games and gamified applications for participatory processes: (i) to illustrate complex urban issues and make the complexity more tangible, (ii) to evoke social learning and capacity building, and (iii) to make the participatory processes 'lighter' and easier to attend.

As to our first point regarding the illustration of complex urban issues, serious games are experienced as suitable formats to illustrate the complexity of urban matters and to make them more tangible. The real-world complexities are then mirrored in the artificial game context. Hence, relations or outcomes of decision-making processes that in the 'real-world' are difficult to experience become visible in the game. Topics such as energy transition, urban planning and urban matters include multiple tiers of policy, a broad variety of actors, conflicting policies, and they touch on politically and societally charged topics: "Perhaps one can say, the more complex the issue is, the more likely the game can achieve something" (VIE-MF) or "Everything can be more accessible via the use of games, certainly mostly the politically and societally loaded topics." (GEN-LA).

Instigating learning and capacity building (Gugerell, Jauschneg, Platzer, & Berger, 2017), communicating and understanding conflicting interests of various stakeholders and actor groups are considered pivotal, and games are seen as being significant tools:

> It was about complexity [of the project, A/N] and to make people, participants realise what the interests of the other participants were and to get to the bottom of these interests….You give people different roles they don't have in real life. (GRO-ES)

The quote sheds light on the importance of games as communication and negotiation environments, where different perspectives and viewpoints can be shared, discussed, deconstructed, and negotiated by the players. Within the game setting "you get people in a situation that they are willing to look differently to this map and so they get away a little bit from their sometimes very small private interests" (GRO-ES) and:

> Trying to explore and to immerse oneself into a topic—and you explore and experience many things, that one should consider. But you also get to know the 'other side', it's—yeah—also a communicative process. (VIE-MF)

It illustrates that games as artificial contexts allow actors to step out of their everyday realities and explore alternative perspectives and possible practices. Hence the second reported benefit is that playing games not only supports visualizing complex planning issues but also instigates processes of social learning and civic capacity building. In the interviews, capacity building was framed as obtaining skills and knowledge of the planning processes and related administrative procedures of the public administration and planning departments: "So people can learn how to get involved in the process and also in projects" (GRO-JKK). But games are particularly valued for triggering, facilitating, and consolidating learning processes:

> Games were used as a consolidation of other learning processes: in a heuristic way, heuristic meaning as a structuring aid for the discovery of certain types of knowledge, or discovery of their own strengths and weaknesses in a particular set. (GEN-VVdS)

The material suggests a strong interest in game components and approaches that support negotiation and deliberation, with a particular focus on collaborative settings:

> You have to collaboratively reach the goal [in the game, A/N]. Thus, the game is very similar to a participatory process….That there is a winner in the game? No, I believe that's not favourable—because it contradicts the participatory idea: I do not want a winner. I do not want that the strongest, quickest or smartest will dominate and prevail. (VIE-MF)

Those multiplayer, collaborative games involve a strong social component, contrary to playing alone or against a computer. Multiplayer games are based on interaction with other people (e.g., debating about different interests, exploring a strategy, praising the achievement of other players), but for games to be appealing and fun, they also need competitive elements that make playing with (not necessarily against) other players challenging (see also Wendel & Konert, 2016):

> Everyone can allocate a point to his favourite measure—and in the end there is somebody winning—therefore it also has a competitive element….That means that all of them endeavour, because sometimes the topics are also a bit ‘dry’. (VIE-JG)

Hence, instigating and supporting different types of collaboration, such as building shared knowledge, resolving conflicts and different interests, motivation and joint goal achievement (Guzzetti, Snyder, Glass, & Gamas, 1993) appear as a major requirement throughout our research. A balance between collaboration and competition appears to be a preferable game setting for media and tools that are used in such participatory approaches.

Finally, the reference towards ‘dry’ topics points towards another important benefit addressed in the interviews: games and gamified applications are expected to make participatory processes lighter and more enjoyable: “The advantage of using games is the low threshold, low design, people buy easily into it, they go along with them. Creates a relaxed, fun atmosphere, that is something that is appreciated by the people.” (GEN-PV). By doing so, they are expected to improve the overall quality of the participatory process by ‘playful deliberation’ (VIE-JG), ‘playful engagement’ (VIE-FM) and by increasing the ‘fun-factor’: “I believe, the fun-factor is crucial, when you think through and work with games. Very often the games are so serious—too serious, that I even think by myself ‘there are fun elements missing’” (VIE-MF) and “because it’s something playful, something where the people get a kick out of it” (VIE-SH). Thus, we see the practical importance of balancing serious games with an equilibrium of serious content and game-fun (Harteveld, 2011; Iten & Petko, 2016; Malone & Lepper, 1987) that results in a joyful game and learning experience (Gugerell et al., 2017). However, it is crucial to stress that though serious games primarily serve ‘non-entertainment’ purposes, they still need to be fun and entertaining to a certain degree to meet both the needs of planning practice and the participants’ expectations.

4.2. Perceived Barriers to the Use of Games and Gamified Applications in Participatory Settings

Both urban planners and process facilitators shared an enthusiasm towards the use of games and gamified applications, with external process facilitators being slightly more open in adopting these new formats than the planners. However, despite the generally positive attitude towards games and gamified applications in participatory urban planning approaches, in all case study cities, similar barriers and challenges seem to impede their regular use. In our research, three main barriers are identified: (i) the modest gaming experience of the facilitators and the planning departments, (ii) a resource scarcity that limits the development of and engagement with such tools, and finally (iii) the fear of reluctant adults to make a fool of themselves.

In all case study cities, the professional experience with games or even gamified environments is modest to limited, with only about a third of the interviewees having previously used games or gamified applications in participatory processes. This limited practical experience reduces the understanding of the potential value of such games in participatory processes. This perception is not only limited to planners and facilitators but is present in senior and high ranking representatives of the public administration:

> Personally, I am not from the gaming generation. So, gaming is kind of alien to me….And unfortunately, most of my colleagues are of my age. I’m almost 50 and I am on the average of the municipality, so we have a very old population. As a consequence, there is not a lot of knowledge about gaming, I would actually say there is not enough knowledge. And I think that most people in our government see gaming, you know, as video gaming, doing stuff, shooting people…crashing cars, stuff like that…but gaming as part of a participation process: I don’t think that many people have ever thought about that. (GRO-JKK)

The modest experience with games also makes it hard to assess and estimate resources needed and expenditure for the development and facilitation of games. “The drive to be efficient and not having a lot of money to get things done, I think prohibits us from taking this step to experiment with [games, N/A]” (GRO-ES). This quote illustrates a possible tension between resource availability, applicability, and repeatability of developed games or game components and the expected benefits on the participatory process itself. Time and budget constraints, paired with the expectation that games require a more elaborate development process, compared to ‘traditional’ methods, are decisive impediments to the development and facilitation of games in planning practice: “it takes a lot of time and thinking to develop a game” (GRO-ES) and “Digital games are really time-consuming to produce” (GEN-LB). This constraint is reinforced by the argument that games are mostly tailored to specific spatial contexts or regulatory frameworks and thus cannot be adapted to other topics or conditions: “Regarding games we are not active at all, because it’s too cumbersome—and for each case it would be necessary to develop something separately or be able to convert it” (VIE-MF) or “Yes, for the game development, you must invest something and then I need the option to use it more often—and it only pays off, if I—ok for every Agenda [Agenda21, A/N] action on the topic I can use it at least ten times” (VIE-SH). There is also a certain ‘mystification’ of the game design process; because of their lack of training and confidence, planners question whether games can actually be suitably designed to achieve their goals: “You can design games in so many different ways to so many different objectives, to include and diffuse so many different kinds of knowledge….It is so flexible in format and this represents as well a challenge because to design in that space something that really works, it is not given" (GEN-PV) and "There is no proof that serious games can produce behaviour changes." (GEN-LB).

Finally, another central concern regarding the broader implementation of games and gamified applications were difficulties with the adult participant group, which makes up the majority of those in the participatory processes that were examined. While games are thought of as being suitable tools to attract participants and to "get people enthusiastic" (GRO-AH) serious concerns are also being voiced regarding their broader use. While games and gamified applications are known to work well in collaborative practices with children, teenagers and young adults, serious concerns regarding their applicability with adults were addressed: "My experience is, that it's [playing games, N/A] fantastic for kids; adults dare only rarely to engage—ok, it depends on the setting. Frankly, I have hesitations, how far you are offering that, or not" (VIE-SH) and "The risks in using games are that people see it as childish, there is always a balance you should make in addressing something playfully, without leaving the impression that what you are doing is mere entertainment." (GEN-LA). Those concerns were shared by about a fifth of the interviewees, who expressed their difficult experiences with sceptical and reluctant adults. Facilitators are concerned that participants would either torpedo, leave or discredit the entire process by questioning its seriousness:

> They [participants, N/A] then often said 'that's utter nonsense' and 'what the heck are you doing here'…and with the adults, once one person left. She said, 'that's childish and immature and I don't participate in such a thing'. That happened on the first evening—and she did not come back afterwards" (VIE-MR)

On the other hand, the research also indicates that if this initial reluctance can be broken, adults will also engage in game activities:

> I think most people were beforehand quite reluctant because they didn't come for a game, but they came for a serious discussion….But what you see is that more participants, in the end, say 'oh! now I understand why he or she is doing that'….Like I said, most people are reluctant because they say 'I am not here to play a game. I am here for serious business' (GRO-ES)

and "When playing the game, at the beginning, participants are always a bit reluctant, they are a bit afraid of using colours, images, saying their opinion etc." (GEN-LH). Other facilitators are more willing to abandon the tool because they are too concerned: "Well, M. and H. were very consequent in that regard, they continued with the group-games, but—well, I would not have continued to play them, I would have given up." (VIE-MR). The material illustrates that adults are considered a difficult age group to engage with via games and gamified activities in participatory processes, due to their expectation that they should be participating in and negotiating in 'serious business'. The issue of this being an unusual format and medium that is regularly associated with entertainment and children, does not align with the fact that the average (video) gamer is 35 years old for men and 44 for women (Entertainment Software Association, 2016). However, it clearly indicates that the concern of the facilitators and the reluctance of this age group must be sufficiently considered in both the game design and the participatory process.

5. Conclusions

Even though games and gamified applications are not the panacea to the longstanding issues of civic participation, they do open up new possibilities for engagement and contribute to the diversification of methods and tools available to the facilitators of these processes. Despite the vivid academic debate on serious games and gamification in various planning contexts (Abt, 1969; Devisch et al., 2016; Gordon & Baldwin-Philippi, 2014; Mayer, 2009; Poplin, 2012; Tan, 2014), our research indicates a notable gap between research and practice. The analysis shows that experimentation with games and gamified applications indeed takes place in planning practice and urban governance, but to a much lesser extent than was initially expected, and it should be noted that there are serious concerns regarding their overall applicability.

Facilitators and planners acknowledge the value and benefits of games, to aid the understanding of complex matters, trigger focus group discussion, and to illustrate and support decision-making processes. Hence our research aligns with scholars such as Gee (2005) and Crookall (2010) who illustrate the value of games for learning and capacity building processes. However, modest experience and knowledge, limited resources, and a lack of adaptability of games for differing occasions, cases, and audiences pose impediments to the broader facilitation and use of games in participatory processes. The insufficient education and lack of training of process facilitators cover a variety of participatory methods (Innes & Booher, 2004). With specific regard to games and gamified applications, facilitators' lack of experience results in an inability to clearly estimate their potential and the ways they can assist the participatory practice. The research illustrates that planners and policy-makers do not make a clear distinction between games and gamified applications and use these terms interchangeably, which leads to a certain fuzziness in the practices they adopt and which often results in either disappointment following their application, or to the exclusion of such tools from the participatory process altogether. Consequently, the conscious identification and selection of digital tools and formats for participatory processes is compromised by this fuzziness, adding another layer of conditioning of the participatory practices by the preconceptions and modestly-informed decisions of facilitators. Hence, capacity building of facilitators regarding the new formats which are available can support the emergence of a culture of experimentation with a range of tools and digital media, including games.

To counter the mentioned lack of financial resources and time, the development of smaller game components and mini-games might be a suitable response towards an efficient use of games under such resource constraints. Mini-games can be advantageous for participatory practices because they are easier to balance between generic (to be adaptable to various occasions and projects) and specific (to the address the particular case and position in the process), combining in one tool the two separate attributes that Gordon et al. (2013) have identified. This quality makes it easier for mini-games to both meet the expectations and to fit in the rather tight budgets of planning practice. Finally, for the development and facilitation of games, adult users need particular attention paid to them: engaging adults in co-design and participatory game design processes might be beneficial to address this user group's reservations, while also contributing to the relevance and local embeddedness of the game or gamified application.

Acknowledgements

The article is a delivery of the research project "Playing with urban complexity. Using co-located serious games to reduce the urban carbon footprint among young adults", funded by JPI Urban Europe.

Conflicts of Interest

The authors declare no conflict of interests.

References

Abdul Jabbar, A., & Felicia, P. (2015). Gameplay engagement and learning in game-based learning: A systematic review. *Review of Educational Research*, *85*(4), 740–779.

Abt, C. (1969). *Serious games*. New York: Viking Press.

Ahmed, A. M., Mehdi, Q. H., Moreton , R., & Elmaghraby, A. (2015). Serious games providing opportunities to empower citizen engagement and participation in e-government services. *Proceedings of CGAMES 2015 USA: 20th International Conference on Computer Games: AI, Animation, Mobile, Interactive Multimedia, Educational and Serious Games* (pp. 138–142). New Jersey: Piscataway.

Belluci, A., Jacucci, G., Kotkavuori, V., Serim, B., Ahmed, I., & Ylirisku, S. (2015). Extreme co-design: Prototyping with and by the user for appropriation of web-connected tags. In P. Díaz, V. Pipek, C. Ardito, C. Jensen, I. Aedo, & A. Boden (Eds.), *End-user development. IS-EUD 2015. Lecture notes in computer science. 9083*. Cham: Springer.

Blockworks. (2015). *Climate hope city*.

Bogost, I. (2011). Persuasive games. *Gamasutra*. Retrieved from www.gamasutra.com/view/feature/6366/persuasive_games_exploitationware.php

Bogost, I. (2014). Why gamification is bullshit. In S. P. Walz & S. Deterding (Eds.), *The gameful world: Approaches, issues, applications* (pp. 65–80). Cambridge, MA: MIT Press.

Boyle, A., Connolly, M. T., Hainey, T., & Boyle, M. J. (2012). Engagement in digital entertainment games: A systematic review. *Computers in Human Behaviour*, *28*, 771–780.

Brown, G., & Chin, S. Y. (2013). Assessing the effectiveness of public participation in neighbourhood planning. *Planning Practice and Research*, *28*(5), 563–588.

Clicks and Links. (2015). *Age of energy*. Amsterdam Smart City, Cityzen.

Conroy, M. M., & Evans-Cowley, J. (2005). Informing and interacting: The use of e-government for citizen participation in planning. *Journal of E-Government*, *1*(3), 73–92.

Crookall, D. (2010). Serious games, debriefing, and simulation/gaming as a discipline. *Simulation & Gaming*, *41*(6), 898–920.

Deltares. (2016). *Port of the future*.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining "gamification". In *MindTrek'11* (pp. 9–15). Tampere: ACM.

Devisch, O., Poplin, A., & Sofronie, S. (2016). The gamification of civic participation: Two experiments in improving the skills of citizens to reflect collectively on spatial issues. *Journal of Urban Technology*, *23*(2), 81–102.

DiSalvo, B., & Morrison, B. (2011). Critique of "gamification" in Khan Academy. In *DiGRA 2011*.

Duke, R. (1975). *Metropolis: The urban systems game*. New York: Gamed Simulations, Inc.

Ekman, J., & Amnå, E. (2012). Political participation and civic engagement: Towards a new typology. *Human Affairs*, *22*(3), 283–300.

Entertainment Software Association. (2016). *Essential facts about the computer and video game industry*. Washington, DC: Entertainment Software Association.

Erhel, S., & Jamet, É. (2013). Digital game-based learning: Impact of instructions and feedback on motivation and learning effectiveness. *Computers and Education*, *67*, 156–167.

Fainstein, S. S. (2000). New directions in planning theory. *Urban Affairs Review*, *35*(4), 451–478.

Fiorina, M. P. (1999). Extreme voices: A dark side of civic engagement. In *civic engagement in American Democracy* (pp. 395–425). Washington, DC: Brookings Institution.

Forrester, J. (1969). *Urban dynamics*. Waltham, MA: Pegasus Communications.

Gee, J. (2005). Why are videogames good for learning? *Spectrum*, *32*, 25–32.

Gläser, J., & Laudel, G. (2010). *Experteninterviews und qualitative Inhaltsanalyse*. VS Verlag für Sozialwissenschaften.

Glick, D. (2012). *Bottom-up urbanism: A survey of temporary use in Europe*. Boston: Hart Howerton Community Fellowship. Retrieved from https://issuu.com/david.t.glick/docs/bottom-up_urbanism_dglick

Gordon, E. (2010). *Participatory Chinatown*. Engagement Lab, Emerson College.

Gordon, E., & Baldwin-Philippi, J. (2014). Playful civic learning: Enabling reflection and lateral trust in game-based public participation. *International Journal of Communication*, *8*, 759–786.

Gordon, E., & Mihailidis, P. (2016). *Civic media: Technology, design, practice*. Cambridge, MA: MIT Press.

Gordon, E., Balwin-Philippi, J., & Balestra, M. (2013). *Why we engage: How theories of human behavior contribute to our understanding of civic engagement in a digital era*. Retrieved from http://cyber.law.harvard.edu/publications/2013/why_we_engage

Grahan, S., & Marvin, S. (2001). *Splintering urbanism: Networked infrastructures, technological mobilities and the urban condition*. London: Routledge.

Gramberger, M. (2001). *Citizens as partners: OECD handbook on information, consultation and public participation in policy-making*. Paris: OECD.

Granic, I., Lobel, A., & Engels, R. C. (2014). The benefits of playing video games. *American Psychologist*, *69*(1), 66–78.

Gugerell, K., Jauschneg, M., Platzer, M., & Berger, M. (2017). Playful participation with urban complexity: Evaluation of the co-located serious game mobility safari in Vienna. In *REAL CORP 2017* (pp. 413–420). Vienna.

Guzzetti, B. J., Snyder, T. E., Glass, G. V., & Gamas, W. S. (1993). Promoting conceptual change in science: A comparative meta-analysis of instructional interventions from reading education and science education. *Reading Research Quarterly*, *28*(2), 116–159.

Handley, D. M., & Howell-Moroney, M. (2010). Ordering stakeholder relationships and citizen participation: Evidence from the community development block grant program. *Public Administration Review*, *70*(4), 601–609.

Harteveld, C. (2011). *Triadic game design: Balancing reality, meaning and play*. London: Springer.

Hay, C. (2007). *Why we hate politics*. Cambridge, MA: Polity Press.

Horelli, L. (2002). A methodology of participatory planning. In R. Bechtel & A. Churchman (Eds.), *Handbook of environmental psychology* (pp. 607–628). New York: John Wiley.

Houghton, K., Miller, E., & Foth, M. (2014). Integrating ICT into the planning process: Impacts, opportunities and challenges. *Australian Planner*, *51*(1), 24–33.

Huotari, K., & Hamari, J. (2011). "Gamification" from the perspective of service marketing. In *CHI 2011*. Vancouver: ACM.

Innes, J. E., & Booher, D. E. (2010). *Planning with complexity: An introduction to collaborative rationality for public policy*. London and New York: Routledge.

Innes, J. E., & Booher, D. E. (2004). Reframing public participation: Strategies for the 21st Century. *Planning Theory and Practice*, *5*(4), 419–436.

Irvin, R., & Stansbury, J. (2004). Citizen participation in decision making: Is it worth the effort? *Public Administration Review*, *64*(1), 55–65.

Iten, N., & Petko, D. (2016). Learning with serious games: Is fun playing the game a predictor of learning success? *British Journal of Educational Technology*, *47*(1), 151–163.

Kelley, T., & Johnston, E. (2011). Discovering the appropriate role of serious games in the design of open governance platforms. *Public Administration Quarterly*, *36*(4), 1–24.

Kleinhans, R., van Ham, M., & Evans-Cowley, J. (2015). Using social media and mobile technologies to foster engagement and self-organisation in participatory urban planning and neighbourhood governance. *Planning Practice & Research*, *30*(3), 237–247.

Krasny, E. (Ed.). (2013). *Hands-on urbanism 1850–2012: The right to green*. Hong Kong: MCCM creations.

Lane, M. B. (2005). Public participation in planning: An intellectual history. *Australian Geographer*, *36*(3), 283–299.

Lerner, J. A. (2014). *Making democracy fun: How game design can empower citizens and transform politics*. Cambridge, MA: The MIT Press.

Lindqvist, J., Cranshaw, J., Wiese, J., Hong, J., & Zimmerman, J. (2011). I'm the Mayor of my house: Examining why people use foursquare—A social-driven location sharing application. In *CHI 2011: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2409–2418). Vancouver: ACM.

Luederitz, C., Schäpke, N., Wiek, A., Lang, D. J., Bergmann, M., Bos, J. J., . . . Westley, F. R. (2016). Learning through evaluation—A tentative evaluative scheme for sustainability transition experiments. *Journal of Cleaner Production*, *169*, 61–76.

Macafee, T., & De Simone, J. (2012). Killing the Bill online? Pathways to young people's protest engagement via social media. *Cyberpsychology, Behavior, and Social Networking*, *15*(11), 579–584.

Malone, T. W., & Lepper, M. R. (1987). Making learning fun: A taxonomy of intrinsic motivations for learning. In R. E. Snow & M. J. Farr (Eds.), *Aptitude, learning, and instruction: Conative and affective process analyses* (pp. 223–253). Hillsdale, NJ: Lawrence Erlbaum.

Marichal, J. (2013). Political Facebook groups: Microactivism and the digital front stage. *First Monday*, *18*(12).

Mayer, I. S. (2009). The gaming of policy and the politics of gaming: A review. *Simulation and Gaming*, *40*(6), 825–862.

Mayring, P. (2015). *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. Beltz.

Medema, W., Furber, A., Adamowski, J., Zhou, Q., & Mayer, I. (2016). Exploring the potential impact of serious games on social learning and stakeholder collaborations for transboundary watershed management of the St. Lawrence River Basin. *Water*, *8*(175).

Parés, M., & March, H. (2013). *Short guides for citizen participation. 3: Guide to evaluating participatory processes*. Barcelona: Department of Governance and Institutional Relations.

Parham, D., Parham, T., & Chang, C. (2011). *Neighborland*. New Orleans.

Poplin, A. (2012). Playful public participation in urban planning: A case study for online serious games. *Computers, Environment and Urban Systems*, *36*(3), 195–206.

Poplin, A. (2014). Digital serious game for urban planning: "B3-Design your Marketplace!". *Environment and Planning B: Planning and Design*, *41*, 493–511.

Prandi, C., Roccetti, M., Salomoni, P., Nisi, V., & Jardim Nunes, N. (2017). Fighting exclusion: A multimedia mobile app with zombies and maps as a medium for civic engagement and design. *Journal Multimedia Tools and Applications*, *76*(4), 4951–4979.

Raphael, C., Bachen, C., Lynn, K.-M., Balwin-Philippi, J., & McKee, K. A. (2010). Games for civic learning: A conceptual framework and agenda for research and design. *Games and Culture*, *5*(2), 199–235.

Robertson, M. (2010). *Hide&Seek: Inventing new kinds of play*. Retrieved www.hideandseek.net/2010/10/06/cant-play-wont-play

Rowe, G., & Frewer, L. J. (2000). Public participation methods: A framework for evaluation. *Science, Technology, & Human Values*, *25*(1), 3–29.

Salen, K., & Zimmerman, E. (2004). *Rules of play. Game design fundamentals*. Cambridge, MA: MIT Press.

Sandercock, L. (1994). Citizen participation: The new conservatism. In W. Sarkissian & D. Perglut (Eds.), *The community participation handbook: Resources for public involvement in the planning process* (2nd ed., pp. 7–16). Murdoch: Institute for Science and Technology Policy.

Scholz, R. W., & Tietje, O. (2002). *Embedded case study methods integrating quantitative and qualitative knowledge*. Thousand Oaks: Sage.

Shaffer, D. W., Squire, K. R., Halverson, R., & Gee, J. P. (2005). *Video games and the future of learning* (WCER Working Paper No. 2005-4). Madison: Wisconsin Center for Education Research, School of Education, University of Wisconsin.

Shipley, R., & Utz, S. (2012). Making it count: A review of the value and techniques for public consultation. *Journal of Planning Literature*, *27*(1), 22–42.

Skocpol, T. (1997). The Tocqueville problem: Civic engagement in American democracy. *Social Science History*, *21*(4), 455–479.

Tan, E. (2014). *Negotiation and design for the self-organizing city*. Delft: TU Delft.

Tolmie, P., Chamberlain, A., & Benford, S. (2014). Designing for reportability: Sustainable gamification, public engagement, and promoting environmental debate. *Personal and Ubiquitous Computing*, *18*(7), 1763–1774.

Tonkens, E. (2014, November 2). De participatiesamenleving inhumaan? *De Humanistish Verbond*. Rerieved from www.humanistischverbond.nl/nieuws/de-participatiesamenleving-inhumaan

Tufekci, Z., & Wilson, C. (2012). Social media and the decision to participate in political protest: Observations from Tahrir Square. *Journal of Communication*, *62*(2), 363–379.

Uitermark, J., & Duyvendak, J. (2008). Citizen participation in a mediated age: Neighbourhood governance in The Netherlands. *International Journal of Urban and Regional Research*, *32*(1), 114–134.

van Bueren, E., Mayer, I., Bots, P., & Seijdel, R. (2007). *The DuBes game*. TU Delft.

Wendel, V., & Konert, J. (2016). Multiplayer serious games. In R. Dörner, S. Göbel, W. Effelsberg, & J. Wiemeyer (Eds.), *Serious games. Foundations, concepts and practice* (pp. 211–241). Springer International Publishing.

Whitton, N. (2011). Encouraging engagement in game-based learning. *International Journal of Game-based Learning*, *1*, 75–84.

Yang, K., & Callahan, K. (2007). Citizen involvement efforts and bureaucratic responsiveness: Participatory values, stakeholder pressures, and administrative practicality. *Public Administration Review*, *67*(7), 249–264. doi:10.1111/j.1540-6210.2007.00711.x

Yin, R. K. (2013). *Case study research: Design and methods* (5th ed.). Thousand Oaks: Sage.

Zichermann , G., & Cunningham, C. (2011). *Gamification by design: Implementing game mechanics in web and mobile apps*. Sebastopol, CA: O'Reilly Media.

About the Authors

Cristina Ampatzidou is a PhD Researcher at the University of Groningen on the topic of gaming and urban complexity, as well as being a founder of Amateur Cities. Her research investigates the affordances of new media for collaborative city making. She has previously worked as a researcher at the University of Amsterdam, and a guest teacher at TU Delft. Cristina has been a collaborator of Play the City! Foundation and the Architecture Film Festival of Rotterdam, and is a regular contributor to architecture and urbanism magazines.

Katharina Gugerell is Assistant Professor in Planning and Spatial Design at the Department of Spatial Planning and Environment, University of Groningen, The Netherlands. Current research topics include Living Labs as experimental forms of governance, games and playful approaches in planning, and power-dimensions in governance networks. She teaches at universities in the Netherlands, Austria, and China.

Teodora Constantinescu is an Architect and Urban Designer with a Master's degree from the Catholic University of Leuven, Belgium. As of 2014, she is a researcher within the Spatial Capacity Building research group at Hasselt University, Belgium. She explores themes such as spatial capacity building, spaces of urban migration, digital social innovation, urban games, and spaces of multicultural micro-economies. Her research topic focuses on the role games play in re-valuing spaces for work.

Oswald Devisch is an Associate Professor in Urban Design at the Faculty of Architecture and Arts, Hasselt University, Belgium. He studied at the Catholic University of Leuven, Belgium and the Bartlett School of Architecture, London, UK. He attained a doctoral degree at the Technical University of Eindhoven, The Netherlands, on spatial simulation models. At Hasselt University, he coordinates the research cluster Spatial Capacity Building exploring themes such as collective learning, casual participation, autonomous transformation processes, and the gamification of participation.

Martina Jauschneg (Dipl.-Ing.) is self-employed directing an office for landscape and open space planning, a Lecturer at the BOKU in Vienna and at the University of Applied Sciences in Salzburg, working in the areas of open space planning, participation and mobility.

Martin Berger (Univ. Prof. Dr.-Ing.) is a Professor of Transport Planning at the Vienna University of Technology. His areas of work include mobility research, municipal and regional transport planning, mobility surveys, and mobility management.

COGITATIO

Urban Planning (ISSN: 2183–7635)
2018, Volume 3, Issue 1, Pages 47–57
DOI: 10.17645/up.v3i1.1299

Article

Essential Means for Urban Computing: Specification of Web-Based Computing Platforms for Urban Planning, a Hitchhiker's Guide

Pirouz Nourian [1,]*, Carlos Martinez-Ortiz [2] and Ken Arroyo Ohori [1]

[1] Faculty of Architecture and Built Environment, TU Delft, 2628 BL Delft, The Netherlands; E-Mails: p.nourian@tudelft.nl (P.N.), g.a.k.arroyoohori@tudelft.nl (K.A.O.)
[2] Netherlands eScience Center, 1098 XG Amsterdam, The Netherlands; E-Mail: c.martinez@esciencecenter.nl

* Corresponding author

Submitted: 30 November 2017 | Accepted: 23 February 2018 | Published: 29 March 2018

Abstract
This article provides an overview of the specifications of web-based computing platforms for urban data analytics and computational urban planning practice. There are currently a variety of tools and platforms that can be used in urban computing practices, including scientific computing languages, interactive web languages, data sharing platforms and still many desktop computing environments, e.g., GIS software applications. We have reviewed a list of technologies considering their potential and applicability in urban planning and urban data analytics. This review is not only based on the technical factors such as capabilities of the programming languages but also the ease of developing and sharing complex data processing workflows. The arena of web-based computing platforms is currently under rapid development and is too volatile to be predictable; therefore, in this article we focus on the specification of the requirements and potentials from an urban planning point of view rather than speculating about the fate of computing platforms or programming languages. The article presents a list of promising computing technologies, a technical specification of the essential data models and operators for geo-spatial data processing, and mathematical models for an ideal urban computing platform.

Keywords
dataflow programming; decision support; planning support; spatial computing; urban computing

Issue
This article is part of the issue "Crowdsourced Data and Social Media for Participatory Urban Planning", edited by Bernd Resch (University of Salzburg, Austria), Peter Zeile (Karlsruhe Institute of Technology, Germany) and Ourania Kounadi (University of Salzburg, Austria).

1. Introduction

In this article we focus on the applications of urban computing in *Smart Cities Planning* practice (as proposed by (Batty et al., 2012)). They suggest that there is a need for a paradigm-shift in urban planning, from focus on the built environment problems to social problems such as deprivation, and their relations to space, spatial distributions and spatial planning. Considering the complexity of cities, they imply that there is a need to develop "a new science of human [spatial] behaviour". This paradigm shift towards developing new [spatial] sciences of cities can be facilitated by the so-called urban computing practices, e.g., by facilitating access to large datasets on human spatial behaviour. This article seeks to illustrate what are the essential means of urban computing practice from a methodological point of view, i.e., computational requirements for 1) developing scientific knowledge in the form of validated analytic/simulation models using spatial data and spatial relations; and 2) informing planning actions using the insight gained from analytic/simulation models on effectiveness of actions.

1.1. What is Urban Computing?

It is difficult, and perhaps even futile, to provide a comprehensive definition of the emerging fields of *Urban Computing* (e.g., as referred to in Kindberg, Chalmers, & Paulos, 2007; Zheng, Capra, Wolfson, & Yang, 2014) and the closely related field of *Urban Informatics* (e.g., as referred to in Foth, Choi, & Satchell, 2011). These two are umbrella terms for describing diverse practices involving geo-spatial data analysis related to cities and citizens. While the former has a technical connotation related to sensing, analysis and actuation technologies (Kindberg et al., 2007), the latter is more focused on the computational social sciences applied to analysis of cities. Without attempting to provide a comprehensive definition, we choose to use the term urban computing with a broader scope to refer to all data-intensive 'computational workflows' that can be used for improving urban planning and urban decision-making by providing the means of data acquisition, analysis and simulation, e.g., to reduce traffic congestion or energy consumption. From a technical point of view, urban computing can involve acquisition, integration, and analysis of (big) data generated by diverse sources such as sensing technologies and large-scale computing infrastructures in the context of urban spaces. The volume, velocity and variety of such data often requires the use of cloud computing infrastructure and software services (Hashem et al., 2015). Urban Computing is applicable in a variety of fields, namely:

- environmental studies (e.g., Shang, Zheng, Tong, Chang, & Yu, 2014; Zheng, Liu, & Hsieh, 2013);
- modelling energy use/generation (e.g., Simão, Densham, & Haklay, 2009);
- transport modelling (e.g., Zheng, Liu, Yuan, & Xie, 2011);
- monitoring health (e.g., Varshney, 2007);
- epidemiology (e.g., Lopez, Gunasekaran, Murugan, Kaur, & Abbas, 2015);
- social informatics (e.g., Foth, Forlano, Satchell, Gibbs, & Donath, 2011; Pires & Crooks, 2017);
- criminology (e.g., Bogomolov et al., 2014); and
- participatory planning (e.g., Robinson & Johnson, 2016; Tenney & Sieber, 2016).

1.2. Why Is Urban Computing needed in Urban Planning?

In Urban Planning, we are often interested in analysing the so-called what-if scenarios using simulations and projections (Batty & Torrens, 2001). Traditionally, the geo-spatial analysis of intervention scenarios, urban plans, and urban data is done by means of Geographic Information Systems (GIS), Planning Support Systems (PSS; see Batty, 2007; Harris & Batty, 1993) and Spatial Decision Support Systems (SDSS). The PSS and SDSS systems are typically stand-alone desktop applications that have a database, a library of computational methods for geo-spatial data processing, and an interface. Despite the technical similarities in using a spatial database, the two categories are different in that the SDSS are geared towards *operational* decision-making whereas the PSS are geared towards *strategic* planning that often involves land-use planning and thus requiring the consideration of land-use transport interactions (the distinction between PSS and SDSS from Geertman & Stillwell, 2009). In these systems, there exist some workflows for spatial analysis of urban data, which do not require new ground-breaking technology. However, the prospect of urban computing is the potentials of the web-based computing platforms for developing a new generation of shareable and editable geo-spatial data processing workflows for informing decisions in urban planning. From urban computing applications listed in Section 1.1, it can be seen that so far urban computing technologies have been mostly applied in the operational and managerial contexts (based on the definition of urban planning actions; Couclelis, 2005). For a wider adoption of urban computing practices in strategic urban planning, urban computing platforms must provide the essential means of analysis and simulation procedures needed in PSS.

Although most of the scholarly works in the area of PSS are focused on land-use change, there are other aspects of urban dynamics that could be modelled computationally; that is to say, the broader discussion is on what changes can be explained, anticipated, and taken into account when making strategic decisions on spatial plans, this broader field of research and development is called Urban Modelling (Batty, 2009). Considering the nature of outcomes of planning processes, (e.g., land-use plans) we can observe that the spatial relations between land-use distributions and a variety of phenomena need to be considered while making strategic planning decisions: for instance, land-use and transport interactions and their effects on energy use in transport (see Keirstead, Jennings, & Sivakumar, 2012) and the effect of land-use distribution on bio-diversity and the use of natural resources (especially water) should ideally be considered when proposing plans. From a pragmatic point of view, however, the adoption of PSS in practice is not high (Geertman & Stillwell, 2009):

> It is disturbing, in fact, to observe the extent to which new computer-based support systems are developed by researchers to the point of adoption but are never implemented in planning practice or policy making. Similarly, there is evidence to indicate that systems which are made operational are not extensively used, after the initial novelty has passed, by those planning organizations for which they have been developed in the first instance. In terms of application, it is possible to point to more failures than successes, i.e., to more cases where systems have not been implemented than examples where they are used routinely. Moreover, many state-of-the-art systems appear to

take a long time to reach the ‘market’ and this is often a process requiring considerable financial resources.

We suggest that the research and development culture of Spatial Planning and Decision Support Systems (SPDSS, terminology of Geertman & Stillwell, 2009) must adopt *open-source and agile* development principles for effective ‘market’ uptake and ensuring the viability of the R&D products (Crowston & Howison, 2005; Hey & Payne, 2015; Pressman & Roger, 2009; von Krogh, 2003). By adopting urban computing practices, utilization of scientific knowledge in planning practice will be eased; because web-based computing platforms facilitate rapid prototyping, development, release, sharing, and test of SPDSS (incorporating a variety of Urban [Analysis/Simulation] Models).

1.3. Problem Statement

Although much can be said about the graphical user interfaces of GIS applications, we do not focus on them; because these interfaces are generally geared towards manual operations. Instead our focus is on the essential means for developing ‘geo-spatial computing workflows’. Workflows can be as simple as routines of sequential actions or more sophisticated procedures with flow-control mechanisms, which are better known as algorithms (see Figures 1 and 2 for workflow examples). There are two types of challenges in using the currently available GIS desktop applications for innovative inter-disciplinary research in Urban Computing applied in Urban Planning (i.e., Design and Development of Web-Based SPDSS):

- Data-Related Challenges:
  - Data-Availability: how easy is it to acquire a relevant dataset?
  - Data-Interoperability: how easy is it to read/write datasets from/to file formats?
  - Data-Mergeability: how easy is it to overlay multiple datasets?
- Workflow-Related Challenges:
  - Workflow Comprehensibility: to what extent is the whole workflow understandable?

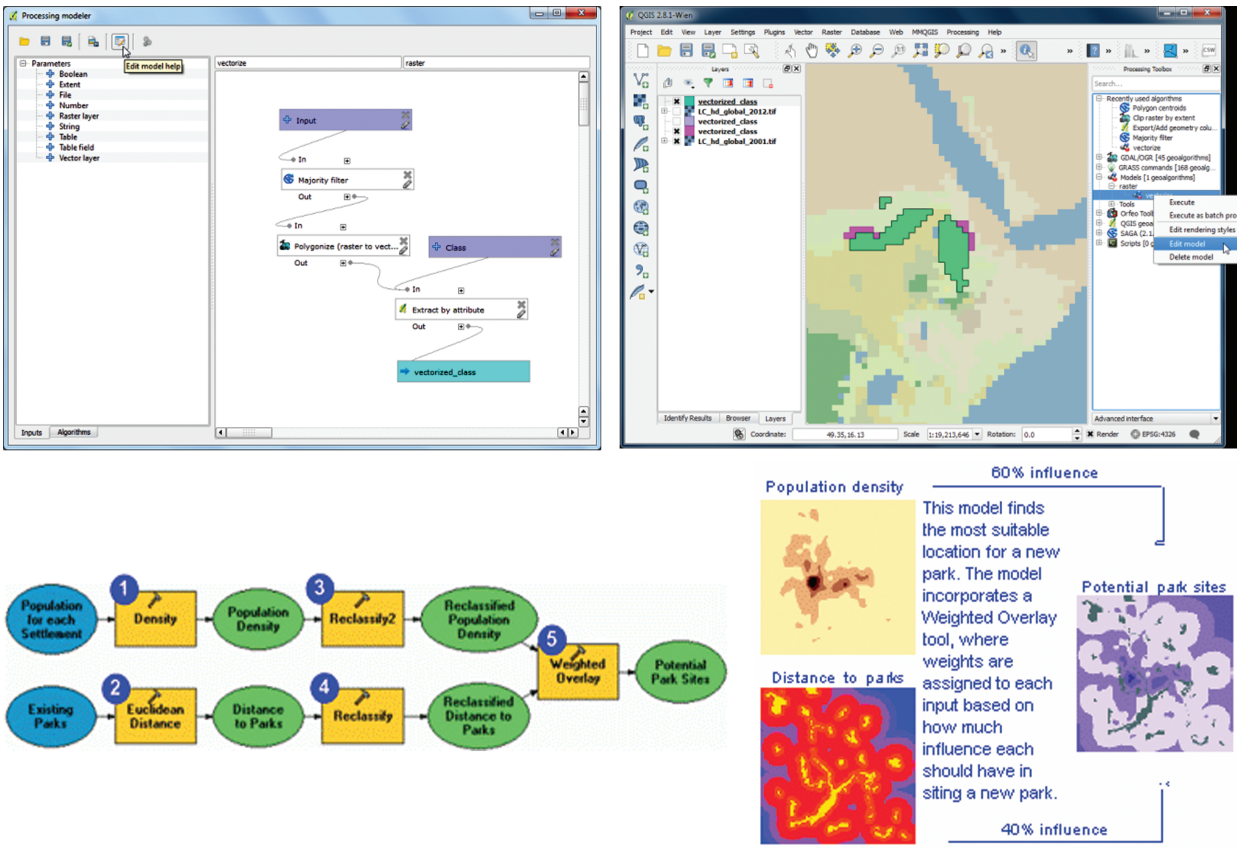


Figure 1. Two examples of geo-spatial data processing workflows from QGIS Processing Modeller[1] (top) and ArcGIS Model Builder[2] (bottom), respectively made for calculating area of water within 25 metres of urban roads (tutorial), and finding suitable locations for urban parks (tutorial).

[1] http://gracilis.carleton.ca/CUOSGwiki/index.php/Automating_Vector_and_Raster_Workflows_using_the_Graphical_Modeler_in_QGIS#Introductions
[2] http://resources.esri.com/help/9.3/ArcGISengine/java/doc/bab90fcc-320b-4b33-902d-a00afd18cfcb.htm

- Workflow Editability: how easy is it to modify the workflow explicitly?
- Workflow Repeatability: how easy is to repeat a certain data processing workflow?
- Workflow Shareability: how easy is it to share a workflow from one system to another?
- Workflow Scalability: how easy is to process large datasets with a workflow?
- Workflow Sustainability: to what extent is the workflow modular and recyclable?

A rather neglected matter about SPDSS is the very social/human process of developing them. These systems can be developed by Research Software Engineers.[3] A typical research software developer is not necessarily a software engineer, but usually a domain-specific researcher who can develop software or computational workflows. A typical research software engineer, often does not have the means of a software vendor to develop a large application with a custom-made GUI. The core of the work of research software development is on developing analytic workflows.[4]

2. What Do We Need for Urban Computing?

We argue that there are three determining factors to consider with regards to 'the suitability of a computing technology for urban computing', i.e., the availability and quality of:

1. Visual Data Flow Programming
2. Spatial Computing Libraries
3. Internet of Things (IoT) APIs[5]

2.1. Visual Dataflow Programming

It is well known that the time spent on research and development is often much more valuable than the computation time. Therefore, we need to consider human interface requirements with regards to the ease of ideation-development-test cycles (prototyping). We propose that using a dataflow programming platform, the user can interact with the platform knowing only a common programming language to edit the nodes (blocks of code) and only a handful of UI manoeuvres to get started; without the problem of learning a sophisticated UI. In processing big data, there are two generic approaches, namely: batch processing and real-time processing (Hashem et al., 2015). Considering the real-time data processing requirement, especially in dealing with managerial and operational planning actions, we can conclude that the Dataflow Programming[6] is an appropriate paradigm for setting up an R&D/prototyping environment (Blackstock & Lea, 2014; Szydlo, Brzoza-Woch, Sendorek, Windak, & Gniady, 2017). Considering that the sustainability and the repeatability of the workflow, it is practical to adopt a modularization and standardization approach to workflow development. Standardization is important for reusability. Specifically, the code-blocks (alias nodes, blocks, or subsystems) of a workflow must input and output data in formats readable for one another. Of course, having a visual overview of the workflow is of high added value, as it makes the workflow as intuitive as a flowchart. The idea of a visual dataflow programming language is to represent the high-level logic of a program/workflow as a graph of nodes, which are blocks of (reusable/shareable) code. The representation of the high-level logic as a graph makes it easy to focus on the complex big-picture for a group of developers working on a workflow. Instead of developing a complete software application with a graphical user interface, a research software engineer can focus on the core of the workflow, model the workflow, test it, share it, and release it as a functional prototype.

If the workflow description language is a (de facto) standard, the intended user does not need to learn a new interface to interact with the workflow. In other words, instead of focusing on optimizing a new software application in terms of its interface and the computational efficiency, more attention can be paid to the effectiveness of the workflow itself. In addition, if the workflow is also cloud-based, then it will be easier to share them and collaborate on-line in real-time.

In short, adopting a visual cloud-based dataflow processing language (and ecosystem) brings about a few advantages:

- Automation of repetitive tasks for data cleansing, validation, etc.;
- Informal and yet sustainable standardization based on common-practices and bottom-up emergence of workflow patterns[7];
- Sharing workflow pattern solutions instead of re-inventing the wheel;
- The possibility of interdisciplinary collaboration;
- Ultimate modularization of workflows based on sharing nodes/blocks of code;
- Agile development-test-release cycles;
- Promotion of Open-Source development practices and therefore rapid progress;
- Ensuring re-usability and repeatability of workflow-based practices such as spatial analyses;
- Saving time by significantly reducing the time and effort in re-inventing interfaces;
- Raising the level of comprehensibility of analytic workflows by providing a glass-box view of the process (as opposed to black-box SPDSS); and
- The possibility of public participation in planning

[3] http://rse.ac.uk/who/
[4] http://www.commonwl.org/
[5] Application Programming Interfaces.
[6] https://stackoverflow.com/questions/461796/dataflow-programming-languages/2035582
[7] http://www.workflowpatterns.com/

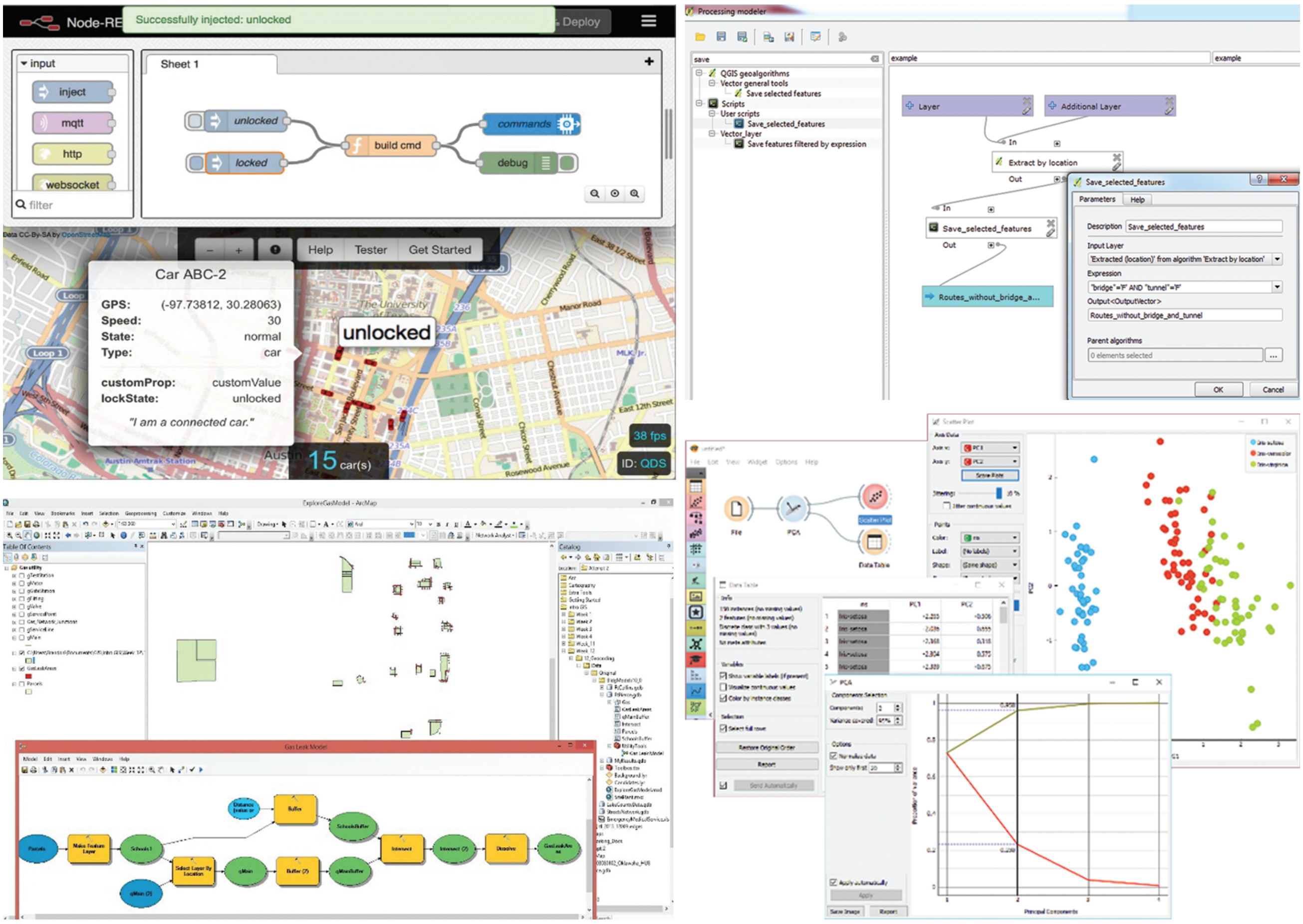


Figure 2. Data processing workflow examples, respectively from top left, clockwise, node-RED, editable by JavaScript (picture from Boyd, 2015), QGIS Graphical Modeller[8], Anaconda Orange3[9], and ArcGIS Model Builder[10], all of which offer Python APIs. The GIS dataflow programming environments make it easy to automate routines, share them, and use standard modules; however, the installation procedures, their domain specific nature and their UI make them much less accessible than the two all-purpose data-flow programming environments shown.

processes by means of rapid development and integration of apps (e.g., using Node-RED[11], a visual data-flow programming tool for wiring together hardware devices, APIs and online services, see Figure 2).

2.2. Spatial Computing Libraries

Here we provide an overview of the requirements of a software application for urban computing; and focus on the specific functionalities that deal with geo-spatial data. Geo-spatial data can be analysed in at least five spatial forms from the most concrete to the most abstract:

- Geographical Data Models: geographically positioned points, lines, polygons, and polyhedrons;
- Geometrical Data Models: points, lines, polygons, and polyhedrons (in local coordinate systems);
- Topological Data Models: vertices, edges, faces, and bodies (algebraic\combinatorial topology);
- Graphical Data Models: objects and links (Graph Theory); and
- Spectral Data Models: eigenvectors and eigenvalues.

The use of the last category of data models is relatively newer than the other types of the models and is used for modelling the dynamics of diffusion flows and Markov Processes in networks (Nourian, 2016; Nourian, Rezvani, Sariyildiz, & van der Hoeven, 2016; Volchenkov & Blanchard, 2007; Wei & Yao, 2014). Performing spectral analyses requires using a computational linear algebra library such as NumPy[12]. Generally, considering the inter-disciplinary nature of urban computing, evident in the breadth and variety of practices mentioned in Section 1.1, we propose that scientific and numerical com-

[8] https://docs.qgis.org/2.8/en/docs/user_manual/processing/modeler.html?highlight=workflow
[9] https://orange.biolab.si/screenshots/
[10] http://pro.arcgis.com/en/pro-app/help/analysis/geoprocessing/modelbuilder/what-is-modelbuilder-.htm
[11] https://nodered.org
[12] http://www.numpy.org/

puting libraries must be available in an ideal platform for urban computing.

In Table 1, we have shown the computational modules required to make spatial analysis and spatial simulation models, which are, in other words, the essential data-models and operations in geo-spatial data processing for urban computing. Central to this schema are the three distinct ways of modelling *space* as:

- **Manifolds**[13] (often approximated as simplicial complexes);
- **Grids** (a.k.a. 2D/3D raster data models, see Zlatanova, Nourian, Gonçalves, & Vo, 2016);
- **Networks** (a.k.a. [directed/weighted] graphs).

In Figure 3, we have categorized the specifically required functionalities for spatial computing as to the previously introduced fields of application of urban computing. There we have shown an overview of exemplary types of analysis or simulation models for planning support workflows, their typical goals and required data models related to the previously listed areas of applications of urban computing.

2.3. IoT APIs

IoT for smart environment is defined by (Gubbi, Buyya, Marusic, & Palaniswami, 2013) as follows:

Table 1. A list of typical goals, required spatial data types, and analytic (mathematical) or simulation (computational) modelling approaches of urban computing.

	Goal	Typically Required Spatial Data Models	3D?	Exemplary/Potentially Applicable Modelling Methodologies
[Land-Use &] Transport Modelling	understanding potentials (accessibility) and predicting the dynamics of mobility [& land-use change]	road network lines, land-polygons, cellular phone network data, GPS trajectories, etc.	possibly beneficial	Discrete-Choice Modelling, Gravity Models, Agent-Based Modelling (ABM), Cellular Automata (CA), Markov Chains, Operations Research
Sociometrics & Econometrics	understanding potentials, and dynamics of social and economic interactions	demographic data attributed to building, block, district, city, or region polygons, crowd-sourced geo-tagged data points, etc.	probably unnecessary	Markov Chains, Markov Chain Monte Carlo (MCMC), Network Centrality, Artificial Intelligence, Statistical Modelling, Predictive Analytics
Criminology & Crime Prevention	understanding potentials, and dynamics of crime in cities	road-networks, demographics attributed to building, & city polygons, geo-tagged (positioned) spatial crime data, etc.	possibly beneficial	Statistical Modelling, Predictive Analytics, Agent-Based Modelling (ABM), Cellular Automata (CA), Markov Chains, Monte Carlo Simulation
Energy Modelling	understanding potentials, and dynamics of energy use and [renewable] energy generation	3D polyhedral models of buildings, point clouds	necessary	Solar Irradiance Simulation (requiring geometric intersections), Computational Fluid Dynamics (CFD, requiring raster and vector fields and differential operators), Monte Carlo Methods
Environmental Modelling	understanding potentials, and dynamics of environmental threats & opportunities (air pollution, noise, vegetation, etc.)	aerial photos, point clouds, vector maps, raster maps	necessary	Analytic Models and Simulation Models (e.g., CA and ABM), Complex System Dynamics, Hydrology, Complex Adaptive Systems

[13] http://mathworld.wolfram.com/Manifold.html

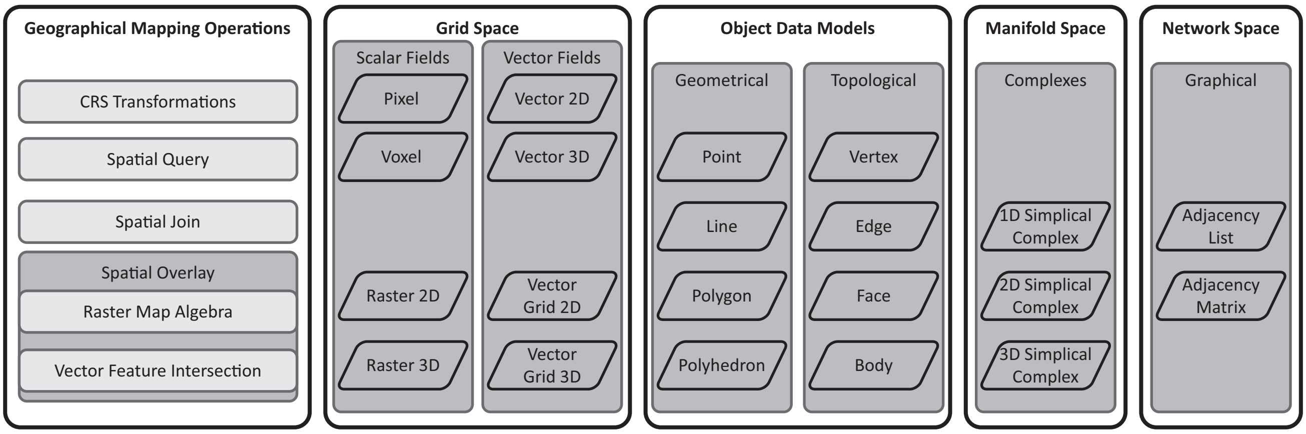


Figure 3. Essential mapping operations and data models required for geo-spatial computing.

> Interconnection of sensing and actuating devices providing the ability to share information across platforms through a unified framework, developing a common operating picture for enabling innovative applications. This is achieved by seamless large scale sensing, data analytics and information representation using cutting edge ubiquitous sensing and cloud computing.

IoT applications can be used for acquisition of data from sensors. They can also be used to directly control some dynamics of cities such as traffic lights. The devices needed for enabling control of physical things are called actuators or actuating devices. The electronic devices that can connect sensors and actuators to internet could be micro-controllers or micro-computers, some of which are open devices popular among amateur enthusiasts such as Arduino[14] and Raspberry Pi[15]. The capabilities of a computing technology for interacting with such devices can be a key factor in making it more pervasive among enthusiast makers and academic software developers, due to the accessibility of such devices in terms of low prices and ease of learning.

Operational planning actions can especially benefit from actuators and sensors in urban environments. For instance, traffic lights can be actuated (controlled) by a controller system connected to many of both sensors and actuators in real-times (thus having a real-time overview of a city) continuously analysing the data coming from sensors sensing the volume of traffic. In other words, IoT devices can facilitate (real-time) operational planning actions. With regards to the IoT potentials for Urban Computing, it is logical to assume that Web-based GIS services (alias web mapping) are necessary for urban computing. In addition, moving all workflows from desktop applications to web-based platforms makes it easier to share (standardized) workflows and collaborate on them. In the next section we focus on the potentials of four programming languages for setting up web-based computational workflows for geo-spatial data analytics and simulations.

3. Promising Technologies for Urban Computing

We have identified a few promising technologies for urban computing, based on Python, Java, JavaScript and R-Spatial languages. From a practical perspective, we consider their potential in terms of ease of prototyping, geo-spatial mapping, 3D visualization, handling big data, and numerical computing (computational linear algebra). From a mathematical/computational point of view, all required models mentioned in Figure 3 can be rather easily developed on top of a robust computational linear algebra library. Apart from numerical capabilities, we argue that for a research software engineer, the visualization and mapping capabilities are essential to consider while making technical choices.

3.1. Python

This programming language is used for example in the Geoda-Web[16], that is the web-based version of CAST[17] with its spatial analysis library PySal[18] seems to be a promising open-source project. Python is the de facto language of open-source development in the field of Geo information science, e.g., in QGIS, Rasterio[19] and Fiona[20]. Python provides a wide range of libraries for numerical and scientific computing such as NumPy, SciPy and Pandas, which facilitates development. Interactive development environments such as IPython (Interactive Python) (Perez & Granger, 2007) and web-based Jupyter notebooks (Shen, 2014) seems to be a promising technology

[14] https://www.arduino.cc/
[15] https://www.raspberrypi.org
[16] http://spatial.uchicago.edu/geoda-web
[17] https://geodacenter.github.io/CAST/
[18] http://pysal.readthedocs.io/en/latest/users/tutorials/dynamics.html
[19] https://github.com/mapbox/rasterio
[20] https://github.com/Toblerity/Fiona

for prototyping and interactive computing. Some universities have started facilitating the use of Jupyter interactive documents as a common means of exchanging reproducible research products, e.g., on JupyterHub[21], NBViewer[22], or SURF-sara (Templon & Bot, 2016) provide hosting and viewing services for sharing Jupyter notebooks. A few options which stand out for simple 3D visualization in Python are: MatPlotLib[23], Mayavi[24] or VisPy[25], while more high-performance applications can be built in OpenGL using PyOpenGL[26]. Web mapping in Python is possible by means of GeoDjango[27].

3.2. Java

This programming language is used for example in a web-GIS for environmental analyses by (Zavala-Romero et al., 2014). The FIWARE platform (Zahariadis et al., 2014) offers an "Application MashUp Generic Enabler", i.e., the WireCloud[28] for visual programming and prototyping web applications. Another flow-based programming environment for Java development supported by Apache Hadoop[29] is NiFi[30]. Java can also provide for interactivity and 3D visualization. The OpenGeoSpatial foundation (aka OSGeo[31]) also provides an open source GIS toolkit for Java called GeoTools[32]. Considering the might of Hadoop for big data analytics and the support of OSGeo Java seems to be a fertile language for urban computing. One option for 3D visualization in Java is JogAmp[33], while a more advanced option is JOGL[34].

3.3. JavaScript

This programming language is used for example in OpenLayers[35] and Carto[36] SaaS (Software as a Service, formerly known as CartoDB[37]) to provide user-friendly WebGIS tools, which can moreover be deployed as desktop applications with tools like Electron[38]. However, neither of them supports explicit workflow development. The other promising JavaScript platform for spatial analysis is MapBox[39], which offers access to the Turf library[40]. Node-RED (Blackstock & Lea, 2014), based on IBM BlueMix (a.k.a. IBM Cloud)[41], seems to be a promising technology in terms of visual programming and the ease of prototyping IoT applications. Node-RED is distributed as part of an open-source software ecosystem called node package manager or NPM[42], that is managed by the Node.js[43] foundation. Interactive visualization in web-browsers is well supported in JavaScript, and arguably more advanced than comparable libraries in Python, thanks to the D3.js library, by Mike Bostock[44] (Bostock, Ogievetsky, & Heer, 2011). In addition to D3 for interactive graphics, there is three.js[45] for WebGL rendering in the browser. Other JavaScript libraries which should not go unnoticed for urban computing are Leaflet[46] (mobile-friendly interactive maps providing access to OSM[47]) and Cesium[48], the latter providing for quality 3D visualization.

3.4. R Spatial

R is a programming language that is part of the R Project for Statistical Computing[49], which includes a complete set of vector algebra operations and functions to create graphics such as plots. The statistical functions in R are much more complete than those available in other languages (e.g., Python). The R Spatial[50] functionality includes the more relevant parts for urban computing, such as representations for raster and vector data, deal-

[21] https://github.com/jupyterhub
[22] https://nbviewer.jupyter.org
[23] https://matplotlib.org/index.html
[24] http://docs.enthought.com/mayavi/mayavi
[25] http://vispy.org/index.html
[26] http://pyopengl.sourceforge.net
[27] https://docs.djangoproject.com/en/dev/ref/contrib/gis/
[28] https://catalogue.fiware.org/enablers/application-mashup-wirecloud
[29] http://hadoop.apache.org
[30] https://hortonworks.com/apache/nifi
[31] http://www.osgeo.org
[32] http://www.geotools.org
[33] http://jogamp.org
[34] http://jogamp.org
[35] http://openlayers.org
[36] https://carto.com/blog/how-to-use-spatial-analysis-in-your-site-planning-process
[37] https://cartodb.github.io/training/intermediate/columbia-sipa.html
[38] https://electronjs.org
[39] https://www.mapbox.com/help/how-analysis-works
[40] http://turfjs.org
[41] https://www.ibm.com/cloud
[42] https://www.npmjs.com
[43] https://nodejs.org/en
[44] https://bl.ocks.org/mbostock
[45] https://threejs.org
[46] http://leafletjs.com
[47] http://www.openstreetmap.org
[48] https://cesiumjs.org
[49] https://www.r-project.org
[50] http://www.rspatial.org

ing with coordinate systems and creating 2D maps. Spatial.ly[51] shows several examples of the more advances visualisation functions in R, including 3D visualisation and animated globes. Shiny[52] is a tool to build web apps with R. There are also other ways in which web sessions of R can be deployed, such as with Rweb[53] and rApache[54]. Similar to Python, Jupyter notebooks can also be used thanks to the IRkernel[55].

4. Conclusion

In response to this question: "What are the essential means for urban computing?", we have provided an overview of specific data models and functionalities required in dealing with geo-spatial data processing (spatial analysis and spatial simulation), referred to as spatial computing in Figure 3 and Table 1, which we deem as the essential means for urban computing. We have considered four programming languages and their promising aspects for urban computing. They all come with their own advantages and shortcomings. It is difficult (and perhaps futile) to point to one of these languages as the most promising language for urban computing. We stress that these technologies are not mutually exclusive, but they can (in some cases) be used in combination with each other. For example, a web-based GIS system could use a Python backend with Flask[56] and a JavaScript frontend with a 3D visualiser based on Cesium, or a processing pipeline could use Python to fetch data from the web using a tool like BeautifulSoup[57], use Java to parse and process the data, use R to do statistical analysis on it, and then visualize the results in a browser using JavaScript. However, it can be said that each of them is stronger in a certain direction, respectively: Java in server-side tools, R Spatial in statistical and mathematical operations, Python in the availability of GIS tools, and JavaScript in IoT and web visualisation. Their respective strengths can be combined by using the best language for each task.

In addition, it is perhaps noteworthy to mention that in the related field of computer-aided design (CAD), there is an active movement towards development of visual programming languages and connecting them together by means of a cloud platform, e.g., Flux[58], initially sponsored by Google[59]. Considering the attractiveness of aligning urban design and urban planning actions, it would be ideal to work in an environment where planners, designers, and research software engineers could all work and share their workflows, for example, a 3D city modelling SaaS such as Möbius[60] (Janssen, Li, & Mohanty, 2016), Tygron[61] or CityZenith[62] could potentially become such a shared development environment.

Acknowledgements

We thank Dr. George Jennings for kindly performing the final proofreading of this article. We are grateful for the meticulous comments of the anonymous reviewers. Ken Arroyo Ohori has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 677312 UMnD).

Conflict of Interests

The authors declare no conflict of interests.

References

Batty, M. (2007). *Planning support systems: progress, predictions, and speculations on the shape of things to come* (UCL Working Papers Series).

Batty, M. (2009). Urban Modelling: Chapter XXX. *International Encyclopedia of Human Geography*. Retrieved from www.elsevierdirect.com/brochures/hugy/overview.html

Batty, M., Axhausen, K. W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., . . . & Portugali, Y. (2012). Smart cities of the future. *The European Physical Journal Special Topics*, *214*(1), 481–518. http://doi.org/10.1140/epjst/e2012-01703-3

Batty, M., & Torrens, P. M. (2001). Modeling complexity: The limits to prediction. *Cybergeo: European Journal of Geography*. http://doi.org/10.4000/cybergeo.1035

Blackstock, M., & Lea, R. (2014). Toward a distributed data flow platform for the Web of Things (Distributed Node-RED). *Proceedings of the 5th International Workshop on Web of Things* (pp. 34–39). http://doi.org/10.1145/2684432.2684439

Bogomolov, A., Lepri, B., Staiano, J., Oliver, N., Pianesi, F., & Pentland, A. (2014). Once upon a crime: Towards crime prediction from demographics and mobile data. In *ICMI'14* (pp. 427–434). Istanbul: ACM. http://doi.org/10.1145/2663204.2663254

Bostock, M., Ogievetsky, V., & Heer, J. (2011). D3 data-

[51] http://spatial.ly/r
[52] https://shiny.rstudio.com
[53] http://pbil.univ-lyon1.fr/Rweb
[54] http://rapache.net
[55] https://irkernel.github.io
[56] http://flask.pocoo.org
[57] https://www.crummy.com/software/BeautifulSoup
[58] https://flux.io/
[59] https://bimandintegrateddesign.com/2014/10/24/googles-bim-busting-app-for-design-and-construction
[60] https://phtj.github.io/mobius
[61] http://www.tygron.com/
[62] http://www.cityzenith.com/smartworld

driven documents. *IEEE Transactions on Visualization and Computer Graphics*, *17*(12), 2301–2309. http://doi.org/10.1109/TVCG.2011.185

Boyd, B. (2015). Build a connected-car IoT app with Geospatial Analytics. Retrieved from https://www.ibm.com/developerworks/library/mo-connectedcar-app/index.html

Couclelis, H. (2005). "Where has the future gone?" Rethinking the role of integrated land-use models in spatial planning. *Environment and Planning A*, *37*(8), 1353–1371. http://doi.org/10.1068/a3785

Crowston, K., & Howison, J. (2005). The social structure of free and open source software development. *First Monday*, *2*(SPEC). http://doi.org/10.5210/fm.v10i2.1207

Foth, M., Choi, J. H., & Satchell, C. (2011). Urban informatics. *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work—CSCW '11*, 1. http://doi.org/10.1145/1958824.1958826

Foth, M., Forlano, L., Satchell, C., Gibbs, M., & Donath, J. (2011). *From social butterfly to engaged citizen: urban informatics, social media, ubiquitous computing, and mobile technology to support citizen engagement*. MIT Press.

Geertman, S., & Stillwell, J. C. H. (2009). *Planning support systems best practice and new methods*. Berlin: Springer.

Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*, *29*(7), 1645–1660.

Harris, B., & Batty, M. (1993). Locational models, geographic information and planning support systems. *Journal of Planning Education and Research*, *12*(3), 184–198. http://doi.org/10.1177/0739456X9301200302

Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Ullah Khan, S. (2015). The rise of "big data" on cloud computing: Review and open research issues. *Information Systems*, *47*, 98–115. http://doi.org/10.1016/j.is.2014.07.006

Hey, T., & Payne, M. C. (2015). Open science decoded. *Nature Physics*, *11*, 367–369. http://doi.org/10.1038/nphys3313

Janssen, P., Li, R., & Mohanty, A. (2016). MÖBIUS, A parametric modeller for the web. In S. Chien, S. Choo, M. A. Schnabel, W. Nakapan, M. J. Kim, & S. Roudavski (Eds.), *Living systems and micro-utopias: Towards continuous designing. Proceedings of the 21st International Conference of the Association for Computer-Aided Architectural Design Research in Asia CAADRIA 2016* (pp. 157–166).

Keirstead, J., Jennings, M., & Sivakumar, A. (2012). A review of urban energy system models: Approaches, challenges and opportunities. *Renewable and Sustainable Energy Reviews*, *16*(6), 3847–3866. http://doi.org/10.1016/j.rser.2012.02.047

Kindberg, T., Chalmers, M., & Paulos, E. (2007). Guest editors' introduction: Urban computing. *IEEE Pervasive Computing*, *6*(3), 18–20. https://doi.org/10.1109/MPRV.2007.57

Lopez, D., Gunasekaran, M., Murugan, B. S., Kaur, H., & Abbas, K. M. (2015). Spatial big data analytics of influenza epidemic in Vellore, India. In *Proceedings of the 2014 IEEE International Conference on Big Data, IEEE Big Data 2014* (pp. 19–24). IEEE. http://doi.org/10.1109/BigData.2014.7004422

Nourian, P. (2016). Configraphics: Graph theoretical methods for design and analysis of spatial configurations. *A+ BE| Architecture and the Built Environment*, *6*(14), 1–348. http://doi.org/10.7480/abe.2016.14.1348

Nourian, P., Rezvani, S., Sariyildiz, I. S., & van der Hoeven, F. D. (2016). Spectral modelling for spatial network analysis. In *Proceedings of the Symposium on Simulation for Architecture and Urban Design (simAUD 2016)*. Retrieved from https://repository.tudelft.nl/islandora/object/uuid:81c02b9c-3ddc-4273-8c2b-7e84c6dc7604

Perez, F., & Granger, B. E. (2007). IPython: A system for interactive scientific computing. *Computing in Science and Engineering*, *9*(3), 21–29. http://doi.org/10.1109/MCSE.2007.53

Pires, B., & Crooks, A. T. (2017). Modeling the emergence of riots: A geosimulation approach. *Computers, Environment and Urban Systems*, *61*, 66–80. http://doi.org/10.1016/j.compenvurbsys.2016.09.003

Pressman, & Roger, S., (2009). *The manifesto for agile software development*, *7*(2). Retrieved from http://nlp.chonbuk.ac.kr/SE/ch05.pdf

Robinson, P. J., & Johnson, P. A. (2016). Civic hackathons: New terrain for local government-citizen interaction? *Urban Planning*, *1*(2), 65–74. http://doi.org/10.17645/up.v1i2.627

Shang, J., Zheng, Y., Tong, W., Chang, E., & Yu, Y. (2014). Inferring gas consumption and pollution emission of vehicles throughout a city. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining: KDD '14* (pp. 1027–1036). http://doi.org/10.1145/2623330.2623653

Shen, H. (2014). Interactive notebooks: Sharing the code. *Nature*, *515*(7525), 151–152. http://doi.org/10.1038/515151a

Simão, A., Densham, P. J., & Haklay, M. (2009). Web-based GIS for collaborative planning and public participation: An application to the strategic planning of wind farm sites. *Journal of Environmental Management*, *90*(6), 2027–2040. http://doi.org/10.1016/j.jenvman.2007.08.032

Szydlo, T., Brzoza-Woch, R., Sendorek, J., Windak, M., & Gniady, C. (2017). Flow-based programming for IoT leveraging fog computing. In *2017 IEEE 26th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)* (pp. 74–79). IEEE. http://doi.org/10.1109/WETICE.2017.17

Templon, J., & Bot, J. (2016). *The Dutch national e-infrastructure*, 13–18. Retrieved from http://pos.sissa.it

Tenney, M., & Sieber, R. (2016). Data-driven participation: Algorithms, cities, citizens, and corporate control. *Urban Planning*, *1*(2), 101–113.

Varshney, U. (2007). Pervasive healthcare and wireless health monitoring. *Mobile Networks and Applications*, *12*(2/3), 113–127. http://doi.org/10.1007/s11036-007-0017-1

Volchenkov, D., & Blanchard, P. (2007). Random walks along the streets and canals in compact cities: Spectral analysis, dynamical modularity, information, and statistical mechanics. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, *75*(2). http://doi.org/10.1103/PhysRevE.75.026104

von Krogh, G. (2003). Open-source software. *MIT Sloan Management Review*, *44*(3), 14–18.

Wei, X., & Yao, X. A. (2014). The random walk value for ranking spatial characteristics in road networks. *Geographical Analysis*, *46*(4), 411–434. http://doi.org/10.1111/gean.12064

Zahariadis, T., Papadakis, A., Alvarez, F., Gonzalez, J., Lopez, F., Facca, F., & Al-Hazmi, Y. (2014). FIWARE Lab: Managing resources and services in a cloud federation supporting future internet applications. In *2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing* (pp. 792–799). IEEE. http://doi.org/10.1109/UCC.2014.129

Zavala-Romero, O., Ahmed, A., Chassignet, E. P., Zavala-Hidalgo, J., Fernández Eguiarte, A., & Meyer-Baese, A. (2014). An open source Java web application to build self-contained web GIS sites. *Environmental Modelling and Software*, *62*, 210–220. http://doi.org/10.1016/j.envsoft.2014.08.029

Zheng, Y., Capra, L., Wolfson, O., & Yang, H. (2014). Urban computing. In *ACM Transactions on Intelligent Systems and Technology*, *5*(3), 1–55. http://doi.org/10.1145/2629592

Zheng, Y., Liu, F., & Hsieh, H.-P. (2013). U-Air. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining: KDD '13* (p. 1436). http://doi.org/10.1145/2487575.2488188

Zheng, Y., Liu, Y., Yuan, J., & Xie, X. (2011). Urban computing with taxicabs. In *Proceedings of the 13th International Conference on Ubiquitous Computing: UbiComp '11* (p. 89). http://doi.org/10.1145/2030112.2030126

Zlatanova, S., Nourian, P., Gonçalves, R., & Vo, A. V. (2016). Towards 3D raster GIS: On developing a raster engine for spatial DBMS. In *ISPRS WG IV/2 Workshop: Global Geospatial Information and High Resolution Global Land Cover/Land Use Mapping*.

About the Authors

Pirouz Nourian (PhD) is an Assistant Professor at the chair of Design Informatics at TU Delft. Pirouz has a PhD in Design Informatics, an MSc in Architecture, and a BSc in Electrical Engineering\Control Engineering. He develops design and research methods and software in the fields of architecture and urbanism. He teaches computational design and procedural 3D modelling in MSc architecture and MSc geomatics at TU Delft. His main areas of research are spatial computation and spatial configuration.

Carlos Martinez-Ortiz (PhD) is an eScience Research Engineer at the Netherlands eScience Center, Amsterdam. Carlos obtained his BSc in computer engineering at La Salle University in Mexico City. At the University of Exeter, he obtained his MSc in applied artificial intelligence and PhD on the topic of shape descriptors for image classification. Carlos has worked on various research projects at the University of Exeter and Plymouth University in collaboration with industrial partners such as C3 Resources Ltd and Met Office.

Ken Arroyo Ohori (PhD) is a Postdoc Researcher at the 3D geoinformation group of the Delft University of Technology in the Netherlands. Ken has a BSc in Computer Science and Technology, an MSc in Geomatics, and a PhD in 3D geoinformation. His current research involves higher-dimensional (4D and higher) data models, data structures and algorithms for Geographic Information Systems. Other research interests include geometric modelling, the validation and repair of geographic data and visualising spatial data.

COGITATIO

Urban Planning (ISSN: 2183–7635)
2018, Volume 3, Issue 1, Pages 58–74
DOI: 10.17645/up.v3i1.1269

Article

Social and Physical Characterization of Urban Contexts: Techniques and Methods for Quantification, Classification and Purposive Sampling

Miguel Serra [1,*], Sophia Psarra [2] and Jamie O'Brien [3]

[1] Research Centre for Territory, Transports and Environment, University of Porto, 4200-465 Porto, Portugal;
E-Mail: mserra@fe.up.pt
[2] Space Syntax Laboratory, Bartlett School of Architecture, University College London, London, NW1 2BX, UK;
E-Mail: s.psarra@ucl.ac.uk
[3] The Bartlett Centre for Advanced Spatial Analysis, University College London, London, W1T 4TJ, UK;
E-Mail: jamie.o'brien@ucl.ac.uk

* Corresponding author

Submitted: 17 November 2017 | Accepted: 26 February 2018 | Published: 29 March 2018

Abstract
Robust quantitative descriptions of the social and physical characteristics of urban contexts are essential for assessing the impacts of urban environments on other, potentially dependent variables. Common methodologies used for that purpose, however, are either coarse or suffer from biasing effects. At the social level, the use of indicators encoded into pre-defined areal units, makes results prone to the Modifiable Areal Unit Problem. At the physical level, the adopted morphological indicators are usually highly aggregated descriptors of urban form. Moreover, there is a lack of explicit methodologies for the purposive sampling of urban contexts with specific combinations of social and physical characteristics, which—we argue—may be more effective than probabilistic sampling, when exploring phenomena as elusive as the effects of urban contextual factors. This article presents a set of GIS-based methods for addressing these issues, based on: a) local indicators of spatial association; b) detailed quantitative morphological descriptions, coupled with unsupervised classification techniques; and c) purposive sampling strategies carried out on the data generated by the proposed context characterization methods (a and b). The methods are illustrated through the characterization of the urban contexts of the 77 state-sector secondary schools in Liverpool, but are generalizable across all categories of urban objects and are independent of the geographical context of implementation.

Keywords
characterization; morphological; purposive sampling; socio-economic; urban context

Issue
This article is part of the issue "Crowdsourced Data and Social Media for Participatory Urban Planning", edited by Bernd Resch (University of Salzburg, Austria), Peter Zeile (Karlsruhe Institute of Technology, Germany) and Ourania Kounadi (University of Salzburg, Austria).

1. Introduction

Urban contextual factors, both social and physical, have impacts on the observed variabilities of a wide range of phenomena as, for example: the distributions of socio-spatial inequalities (Rae, 2012), the spatial incidences of public health problems such as obesity (Townshend & Lake, 2009) and mental health (Cutrona, Wallace, & Wesner, 2006; Miles, Coutts, & Mohamadi, 2011), local differences in patterns of physical activity (Timperio et al., 2010) and mobility (Crane, 2000), or still the spatial distributions of crime occurrences (Charron, 2009). In more general terms, the whole body of literature in the field of 'neighbourhood effects', departs from the hypothesis that local urban contexts have significant impacts on the life of residents, and seeks to assess such hypothesis (van Ham & Manley, 2012).

Nevertheless, any research aiming at identifying potential links between urban contextual factors and other variables of interest, must necessarily face at least two initial problems. The first is how to quantify urban contexts at both social and physical levels. The methodologies currently employed for those purposes have several limitations, and the need for robust quantitative methods has been acknowledged by several authors (Cummins, Macintyre, Davidson, & Ellaway, 2005; Gambaro, Joshi, Lupton, Fenton, & Lennon, 2016; Lupton, 2003). The second problem concerns the criteria for generating context-informed samples of urban areas. As van Ham and Manley (2012) note, quantitative studies using large randomized probability samples have been shown to be far less effective than qualitative studies (i.e., focusing on the experiences and perceptions of residents), in detecting contextual or neighbourhood effects. Qualitative research, however, due to its laborious inquiry processes, demands sampling strategies aimed at creating small yet information-rich samples; that is, purposively selected samples. Purposive sampling strategies are quite different from those of probabilistic sampling, seeking not generalization or randomness, but the well-informed selection of very specific cases, capable of maximizing the chances of observing phenomena of interest. They are also less well-known and understood than probabilistic sampling strategies, even though they are more suitable in certain circumstances (e.g., when the studied population is small) and more effective when used in qualitative research (Patton, 1990).

This article develops a unified methodological framework for the quantification, measurement and sampling of urban contexts, based on both social and physical characteristics, using GIS. The framework was devised within the scope of a community-based, participative research project "Visualising Inequality in Community Networks to Enhance Participatory Planning" (O'Brien, Garcia Vélez, & Austwick, 2017; O'Brien et al., 2016), supported by Leverhulme Trust Research Project Grant, where it was used for characterizing the socio-economic and morphological contexts of all state-sector secondary schools in the Liverpool City Region (Merseyside, UK). However, the framework was designed from the outset so as to remain applicable in any research exercise demanding quantitative contextual characterizations of a given class of urban features (e.g., urban neighbourhoods, the surroundings of institutional buildings or designated public spaces). These quantitative characterizations may be then used to support the generation of purposive samples of local urban areas, informed by a rigorous and detailed characterization of their social and physical differences.

The objective of this article is, therefore, methodological. It aims at contributing to the fields of urban and neighbourhood studies, by providing enhanced context characterization and sampling tools, with a particular focus on purposive sampling strategies. Such tools can be used for several research purposes, namely: a) supporting qualitative, participatory urban research designs, by providing quantitative contextual characterizations of the studied areas, against which qualitative findings may be assessed and interpreted; b) for supporting studies on the complex interactions between social phenomena and the built environment, by enabling the purposive sampling of urban areas with specific combinations of social and physical characteristics; and c) generating samples of local urban areas controlling for their social and physical characteristics, in order to avoid potential confounding effects on the study of other variables of interest.

The article is organized as follows: in Section 2 we discuss the limitations of the current urban contextual characterization methods and how to overcome or mitigate such limitations. We then present a unified methodological framework for characterizing and sampling urban contexts. We end this section by briefly describing the abovementioned research project, as a background to illustrating the application of the methodology. The application of methods is discussed in the following three sections, which constitute the main body of the article. We conclude by summarizing the outputs of the proposed methods.

2. Methodological Framework

Quantitative characterizations of urban contexts are usually carried out at two different levels: social and physical. At the social level, such characterizations usually rely on statistical indicators, spanning several socio-economic and demographic dimensions, commonly aggregated into individual administrative divisions of varying geographies (e.g., census tracts). Although of obvious convenience, due to the wide availability of census data and the pre-defined nature of administrative boundaries, such approach raises several methodological issues (Caughy, Leonard, Beron, & Murdoch, 2013; Kim, Ali, Sur, Khatib, & Wierzba, 2012; Lebel, Pampalon, & Villeneuve, 2007).

Firstly, there is no guarantee that the boundaries of extant administrative geographies will indeed correspond to meaningful spatial or social units of analysis, within the context of each study. However, given the 'off the shelf' availability of administrative boundaries, researchers often use those whose average size better approximates their research objectives. Secondly, because census data are aggregated into administrative units of varying sizes and boundaries, individual units' attributes are prone to be biased by the Modifiable Areal Unit Problem (MAUP; Openshaw, 1983), both through its scale effects (i.e., sensitivity to levels of aggregation) and zoning effects (i.e., sensitivity to the shapes of aggregation units). And thirdly, administrative units are usually characterized and sampled accordingly only to their individual attributes, without regard to their wider spatial embedment (i.e., not taking into account their neighbours' characteristics). However, because individual attributes may be biased by the MAUP, their use as only characterization criterion may not be the best option.

At the physical level, urban contexts are commonly described through broad morphological indicators (e.g., residential density, functional diversity or the total area of green spaces). These indicators are measured and aggregated also at the level of pre-defined administrative units (Charron, 2009; Inoue, Stickley, Yazawa, & Shirai, 2016; MacDonald, Wise, & Harris, 2008; Miles et al., 2011; Townshend & Lake, 2009). Besides being also prone to MAUP (due to the zoning effect), such indicators are rather coarse descriptions of the built environment and may not be sufficiently detailed for detecting potential statistical associations between different urban morphologies and other variables. On the other hand, urban typomorphologies (Vernez-Moudon, 1994) can be of great interest for characterizing physical urban contexts, because they allow identifying comparable and/or contrasting built environments. But urban typomorphologies are commonly identified visually and described only through semantic and graphical means. Constructed in this way, they are difficult to generalise and use outside of their original observation setting. Nevertheless, there is also today a growing body of research dedicated to quantitative methods for the detailed description and classification of urban form, using geocomputation and algorithmic classification methods; see, for example, the work of Gil, Beirao, Montenegro and Duarte (2011) and of Hamaina, Leduc and Moreau (2012). However, to the best of our knowledge, these algorithmic methods have never been applied in neighbourhood or community-based studies, requiring the support of urban physical characterizations.

These quantification shortcomings can affect the identification and sampling of relevant cases, among the variability of social and physical urban contexts. An important decision in studies of local urban communities or neighbourhoods, particularly in the case of qualitative research, is the selection of the cases to be studied. In qualitative research, as opposed to purely quantitative research, the generation of samples is often done purposefully, i.e., in a non-random manner, identifying information-rich cases, in the light of the specific phenomena under investigation (Patton, 1990). Such samples should not include too many cases (for logistical and financial reasons) and should meet defined conditions determined by the phenomena under study or by the question being asked. However, if the quantitative methods adopted for characterizing specific cases are biased in the first place, purposive sample generation based on their results may obviously be jeopardized.

The methodological framework proposed in this article (Figure 1) tries to overcome the abovementioned problems. The framework is divided into two analytical tracks—concerning social and physical urban contexts—which are organised into three steps: 'data preparation', 'quantification' and 'characterization'. The data produced by the two tracks are joined at the end, supporting the generation of purposive samples of local urban areas based on their contextual characteristics.

At the social level and in order to avoid or mitigate the effects of the MAUP, we propose to change focus from the specific values of the variables at each spatial unit (which is the level at which MAUP occurs) to another type of quantitative property, namely the degree of local spatial autocorrelation (LSA) of each spatial unit (regarding the variable under study). This approach has two advantages. Firstly, it mitigates the zoning effects of the MAUP because only the degree of LSA is taken into account for characterizing each spatial unit, to which is associated a probability of it being observed simply by chance. Thus, the selection of spatial units based on the significance of their degree of LSA on a given socio-economic variable, ensures that the unit selected is actually embedded in a (non-random) spatial cluster of similar values. Secondly, it also avoids the scale effects of the MAUP, because LSA may be assessed across several spatial scales, allowing the selection of units that show consistent behaviours across scales. In order to characterize spatial units by their degree of spatial autocorrelation, we use Local Indicators of Spatial Association (LISA; Anselin, 1995), which are geo-statistical methods devised for that purpose. LISA methods and their application are detailed in Section 3.

Regarding the characterization of physical contexts, we move from the coarse descriptions of urban form mentioned previously, to more discriminant methods of quantitative morphological characterization. In the field of urban morphology there has been recently increasing interest in the development of quantitative methods for measuring and classifying urban forms (Barthelemy, 2015; Dibble et al., 2017; Gil et al., 2011; Hamaina et al., 2012; Marshall, 2005; Pont & Haupt, 2010; Serra, Gil, & Pinho, 2016), using available vector datasets of street networks and building footprints, analysing their morphological information through geocomputation and subjecting it to unsupervised classification algorithms. This algorithmic approach produces consistent and quantitatively defined morphological classifications, which are automatically derived only from the morphological data, providing objective criteria for accurately describing similarities and differences in local urban morphologies.

The proposed method adopts a number of morphometric indicators, describing three dimensions of urban form: the network of open space, the geometry of urban blocks and that of building footprints. These indicators are quantified in GIS, using open source vector datasets for each study area. The resulting morphological data are subsequently subjected to unsupervised hierarchical classification, in order to reduce their variability to a manageable number of clusters, representing actual and measurable morphological cleavages between the studied areas. These methods are described in Section 4.

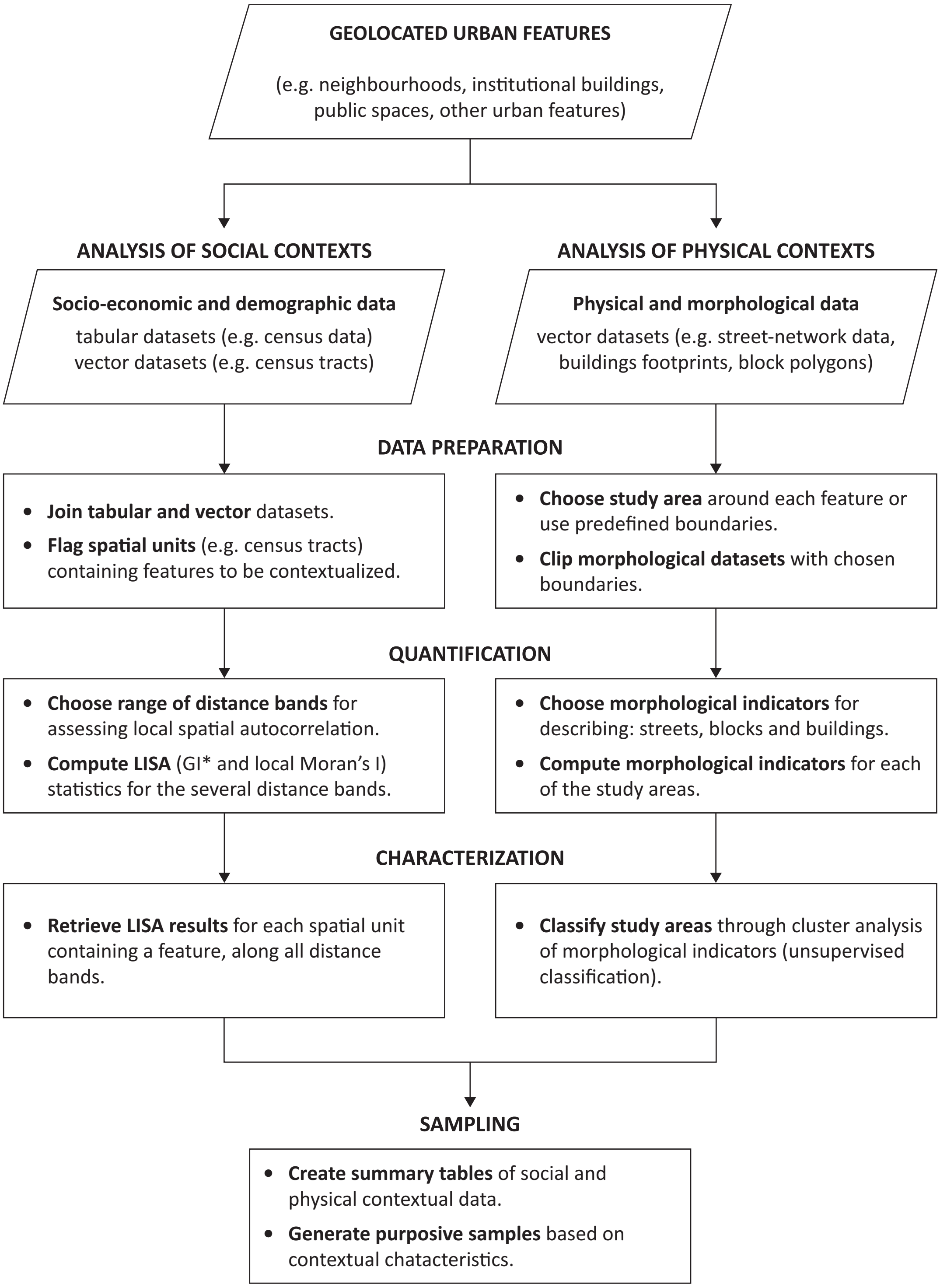


Figure 1. Proposed methodological framework, divided into two analysis tracks, each of which is organised into three stages and a synthesis in order to derive purposive samples.

Finally, the results of both characterization methods are summarized through simple data visualization schemes (i.e., summary tables), allowing for the quick identification of relevant cases, according to several purposive sampling strategies (Patton, 1990), aimed at answering specific research questions. These sampling strategies and the corresponding samples are described in detail in Section 5.

We end this section by providing a succinct description of the research project "Visualizing Inequalities in Community Networks", in the context of which the proposed methodological framework was developed, and whose data we use here for illustration purposes. The project consisted of community-based, qualitative research, in part carried out in Liverpool, Merseyside, UK (Figure 2). The main objective was to gain an understand-

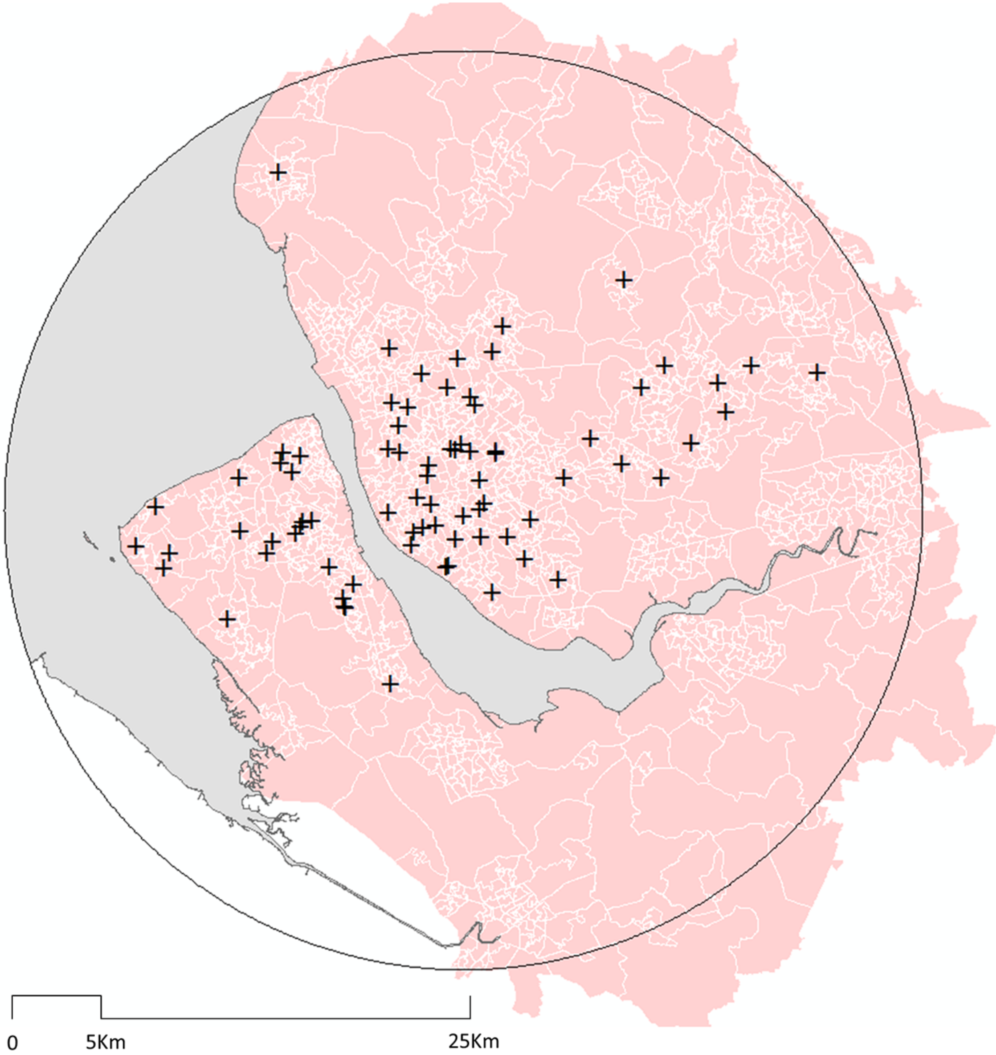


Figure 2. Distribution of the 77 secondary schools in the study area, over LSOA geography.

ing of how people make use of spatial assets in their vicinities, in order to conceptualize their local community formations. In addition, the project sought to understand how the characteristics of the built environment enabled or hindered individual community conceptualizations, and how this might vary across different social and physical urban contexts.

The study focused on the 77 state-sector secondary schools within Liverpool City Region (Figure 1), a region which presents among the UK's widest socio-economic inequalities (LCC, 2015). From the 77 schools, a small sample (n = 16) was chosen (23% of the total) based on the following criteria (by order of relevance): a) the responsiveness of secondary school teachers to our invitation to participate in the study; and b) the inclusion of schools with contrasting social and physical urban contexts. Qualitative data for the project's research purposes, described in detail elsewhere (O'Brien et al., 2017; O'Brien et al., 2016), was gathered through participatory workshops carried out in the 16 selected schools, involving 246 secondary school-age children, aged from 11 to 19 years.

The methods described in this article had a twofold purpose. Firstly, to provide quantitative information of the social and physical urban contexts of all 77 schools, so that that responded positively to the invitation could be evaluated regarding criterion b), above. Secondly, to use the quantitative data as benchmark against which the qualitative data gathered through participatory research activities could be interpreted.

3. Characterizing Urban Socio-Economic Contexts

In order to characterize the socio-economic contexts of all 77 Merseyside secondary schools, we started from a convenience definition of boundaries, namely those of the lower super output areas (LSOAs) where each school is located (Figure 2). LSOAs are geo-located units devised by the UK's Office of National Statistics (ONS) to represent population aggregations by place of residence of around 1500 inhabitants. LSOAs do not represent any meaningful definition of 'neighbourhoods' or of 'urban communities', but they do allow for relatively stable local area analyses because they are designed to have similar population sizes and be as socially homogenous as possible (ONS, 2011).

We load all LSOAs into a GIS within and intersected by a 25 km radius circular boundary, centered on the Liverpool City Region polygonal centroid (Figure 1). As an indicator of their socio-economic composition, we associate each LSOA with its corresponding score in the 'Income Deprivation' domain of the Index of Multiple Deprivation (IMD), the official measure of relative deprivation in England (DCLG, 2015). IMD is a composite index, constructed by weighting indicators for seven domains of deprivation, of which 'Income' and 'Employment' carry the greatest weight (22.5% each). We use the 'Income' domain scores instead of IMD scores, because they are meaningful and interpretable (corresponding to the percentage of the income-deprived population in each LSOA). In contrast, IMD scores are highly transformed and not comparable (IMD scores should be ranked or classified in quantiles) (DCLG, 2015). IMD 'Income' scores, as well as LSOAs boundaries, are provided as open datasets by the ONS.

As previously explained, we are specifically interested in the spatial embeddedness of each LSOA within potentially larger geographical patterns of income deprivation or lack thereof. For this purpose, we use a set of spatial statistics based on the concept of spatial autocorrelation, known under the broad designation of LISA (Anselin, 1995). LISA methods determine the degree to which a geographical feature (e.g., a given LSOA) has a particularly high or low score, according to the attribute itself *and* to the location of the feature in question. In this way, we can evaluate the degree to which a school located within a high-, medium- or low-deprivation LSOA is also located within a larger area of relative deprivation or affluence. Moreover, spatial embeddedness may be assessed for neighbourhoods of varying sizes around each feature. As argued in Section 2, shifting the focus from individual scores to the LSA of those scores, mitigates both the zoning and scale effects of the MAUP.

We apply two LISA methods: the Gi* statistic (Ord & Getis, 1995) and the Local Moran's I statistic (Anselin, 1995), also known as 'hot spot analysis' and 'cluster and outlier analysis', after their respective implementations in ArcGIS 10 (ESRI, 2011). These methods allow comparing each LSOA's 'Income' score with those of its neighbours. When those scores are similarly high or low (i.e., when there is local spatial autocorrelation between a given feature and its neighbours) and when that similitude attains a given degree of statistical significance (i.e., a low probability of occurring by chance), both methods retrieve a signal of spatial clustering (i.e., a significantly high or low Z-score, associated with a certain p-value). Both methods also require the choosing of a given distance band in order to define which neighbouring features are included in the calculations.

The main difference between the two methods concerns the inclusion of the value of the feature under analysis in the calculations of the local mean. In the case of hot/cold spot analysis (Gi*), the local mean is calculated by taking the values of the feature under analysis and those of the features within its neighbourhood; this local mean is then compared to the global mean (i.e., of the entire study area). If significantly different (i.e., yielding a large positive or negative Z-score), the feature is categorized as being part of a hot or cold spot. The output, for each distance band, is therefore a set of hot and/or cold spots, with varying confidence levels (CL, 90%, 95% and 99%) and/or a set of areas without significant clustering (i.e., CL < 90%) of both high or low scores.

In cluster and outlier analysis (local Moran's I) the process is similar, but only the values of the neighbouring features, and not the value of the feature under analysis, are considered. Again, this local mean is compared to the global one, in order to ascertain if they differ significantly. However, the resulting value of the statistic (I) indicates if the feature under consideration also differs from the local mean or not. If the value of I is positive, the feature under analysis has neighbouring features with similar values and is therefore part of a cluster (of high or low values). If the value of I is negative, the feature under analysis has a value that is dissimilar from the local mean and is therefore a local outlier.

We use both LISA methods simultaneously, because they produce slightly different, but complementary outputs. Analysis was run at a range of increasing fixed-distance bands, namely 0.5 km, 1 km, 1.5 km, 2 km, 2.5 km and 3 km. This approach allows us to study the clustering of 'Income' scores across several spatial scales, without the need to define some fixed, discretionary size for the neighbourhood of each feature. The output maps of both methods (Figures 3 and 4) show similar patterns, with high and low 'Income' scores clustering very clearly, revealing the existence of strong socio-economic cleavages within the study area. Central urban areas tend to display high values of deprivation, while peripheral urban centres and rural areas show concentrations of low deprivation. Deprived areas show significant clustering immediately at 0.5 km, whereas non-deprived areas start to cluster at larger scales (from 1 km on), reflecting the different sizes of central and peripheral LSOAs (deprivation is predominant in smaller and more central LSOAs).

Figure 5 shows the socio-economic contexts of each of the 77 schools. For each LSOA containing a secondary school (the columns, in Figure 5), we record its 'Income' score and the rank of that score, regarding the set of 77 secondary schools. We recall that low 'Income' scores mean low deprivation, thus the LSOA ranked first is also the most deprived. Furthermore, we also record the status of each LSOA regarding the spatial clustering of 'Income' scores, as categorized by the two LISA methods across the distance bands mentioned before.

Regarding the results of hot/cold spots analysis (Gi*) and for each distance band (the lines, in Figure 5), each LSOA containing a school may: 1) be part of an area with non-significant clustering of 'Income' scores (white cells); 2) be part of a cold spot of 'Income' scores (blue cells: light blue 90% CL, medium blue 95% CL and dark blue

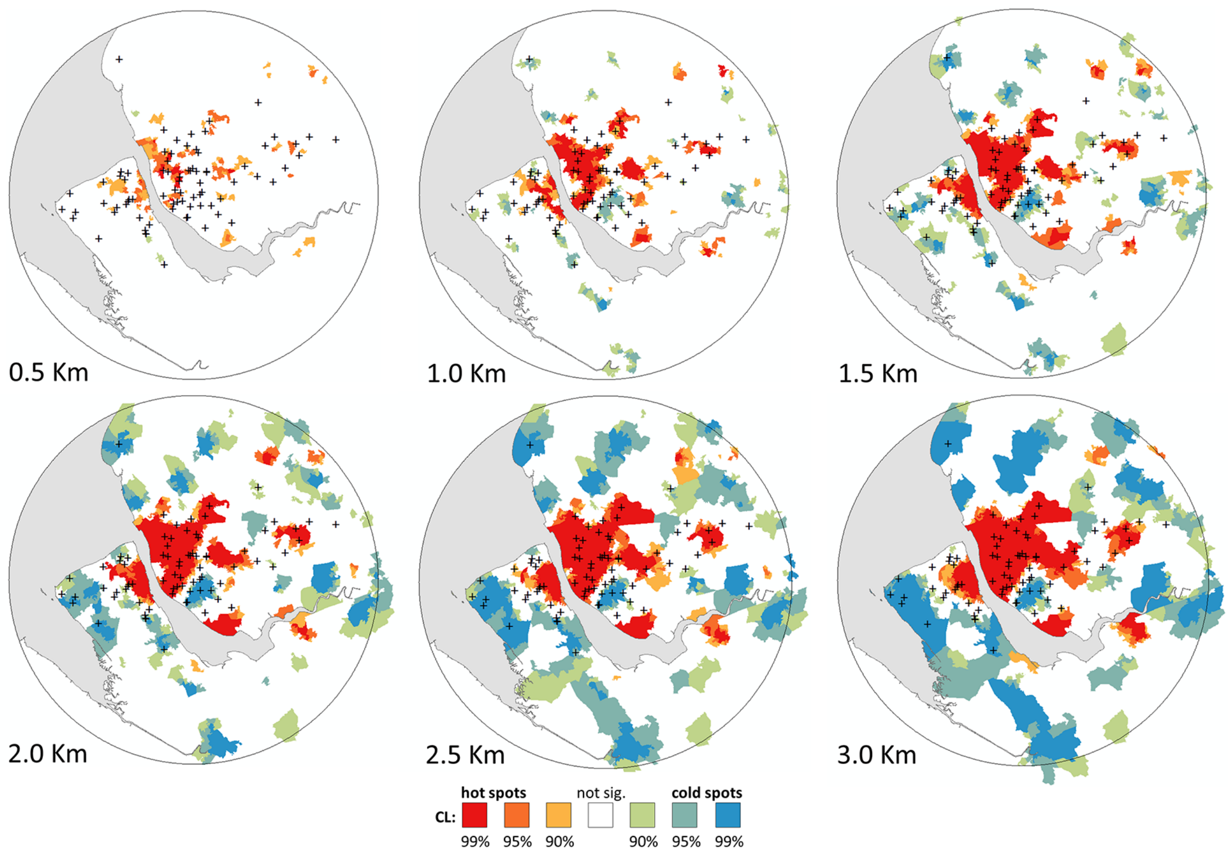


Figure 3. Hot/cold spot analysis of Income scores for the study area.

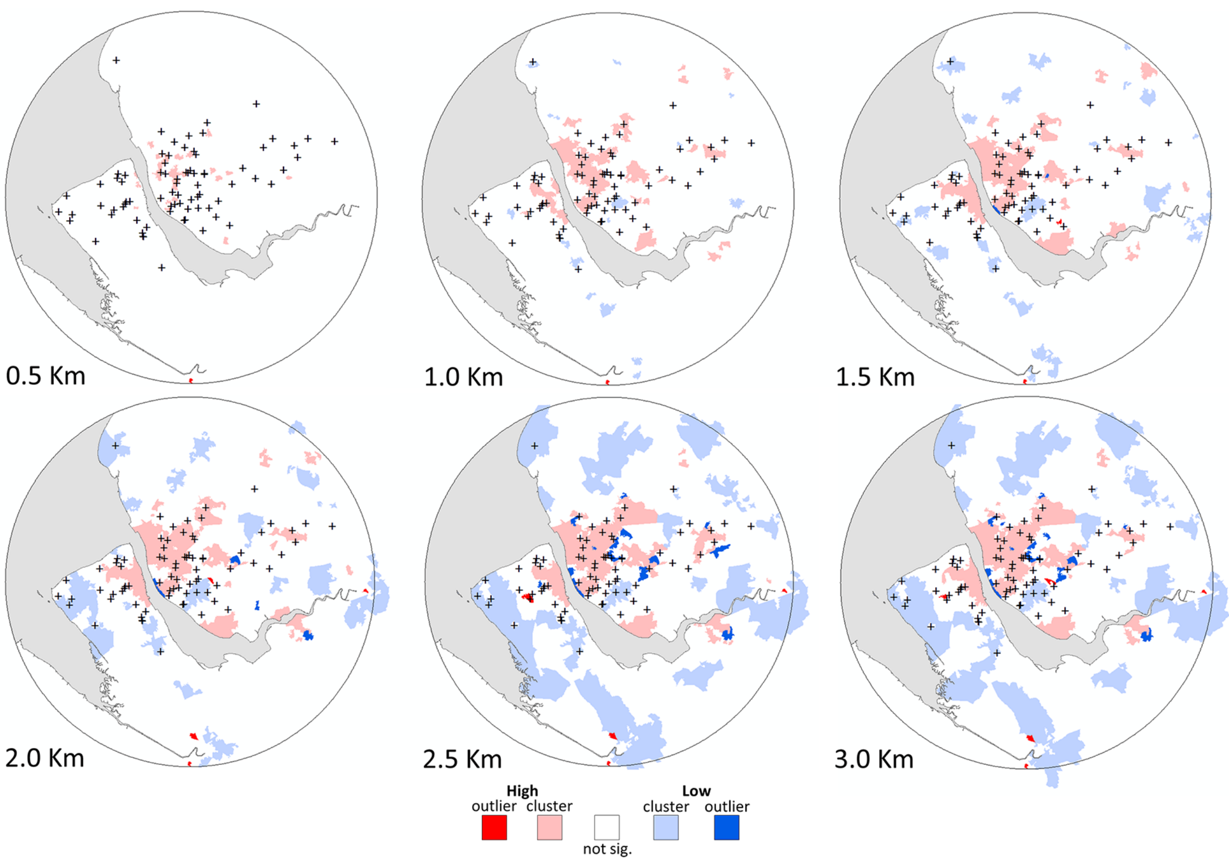


Figure 4. Cluster/outlier analysis of Income scores for the study area.

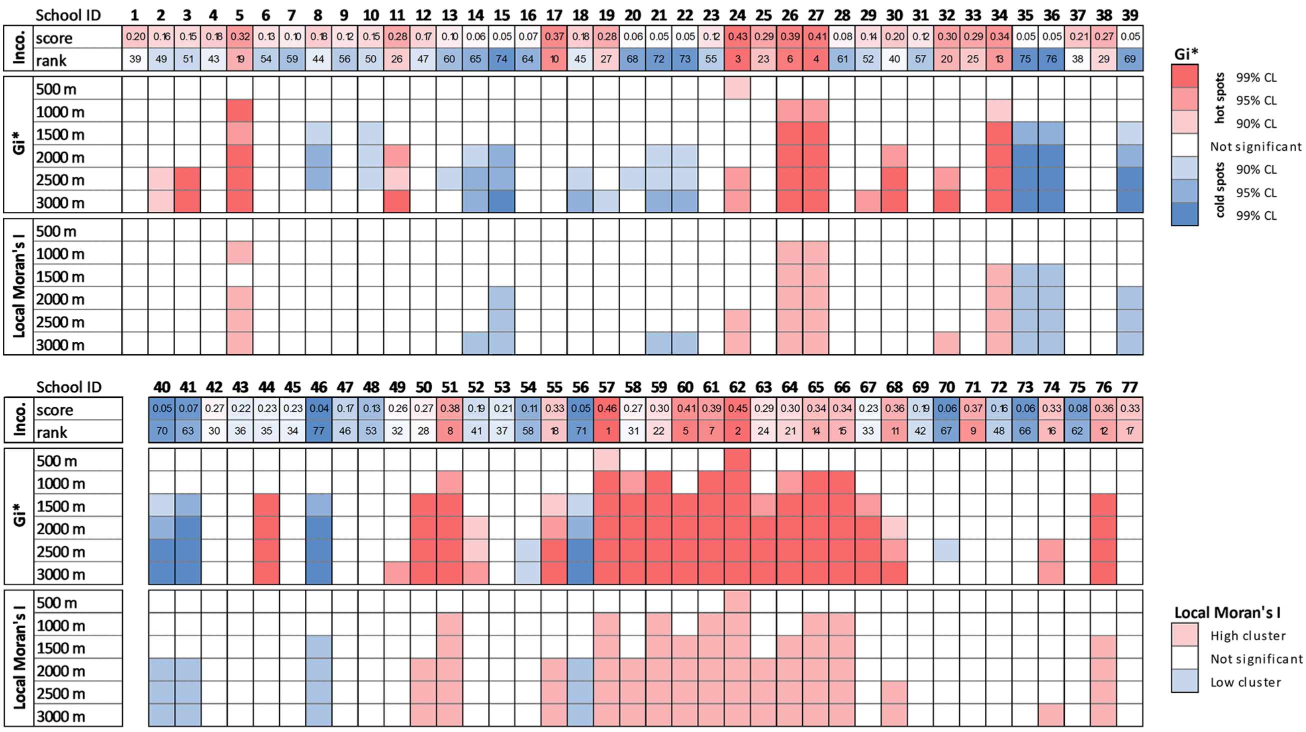


Figure 5. Socio-economic contexts of the 77 secondary schools.

99% CL); and 3) be part of a hot spot of 'Income' scores (red cells: light red 90% CL, medium red 95% CL and dark red 99% CL).

Similarly, regarding the results of cluster/outlier analysis (local Moran's I) and for each distance band, each LSOA containing a school may: 1) be part of an area with non-significant clustering of 'Income' scores (white cells); 2) be part of a cluster of low 'Income' scores (light blue cells, 95% CL); 3) be part of a cluster of high 'Income' scores (light red cells, 95% CL). In the maps of Figure 4 there are several LSOAs that are spatial outliers, but none of the 77 schools is located in these.

Figure 5 shows that there is indeed strong variability in the socio-economic contexts of the secondary schools in the Merseyside. In a number of cases, the school's socio-economic context does not show any significant clustering of either high or low 'Income' scores (e.g., cases 1, 4 or 6) and could be deemed uncharacteristic study targets. Nevertheless, some of these cases do correspond to LSOA that have either high individual scores (e.g., cases 17 and 71) or low individual scores (e.g., cases 7 and 16), in spite of not being part of spatial clusters of neither high nor low scores. We also observe a number of schools within areas where there is a very consistent clustering of high (e.g., cases 5, 27 or 34) and of low (e.g., cases 35, 39 or 46) scores, at almost all spatial scales. These schools are deeply embedded into consistent areas of high and low deprivation and would, therefore, constitute good study targets.

We also noted cases where the categorization produced by the two LISA methods is not consistent. For example, cases 8, 11 or 44, are part of cold/hot spots as defined by Gi*, but not of clusters of low/high scores, as defined by local Moran's I. These differences happen because of the different calculations used in the two methods, but they also speak to a lesser consistency in the characterization of these cases. There are therefore advantages in employing both methods, because they allow assessing the consistency of results and the eventual rejection of inconsistent cases.

4. Characterizing Urban Morphological Contexts

In order to characterize the urban morphological contexts of the 77 schools, we start by defining circular areas of 1 km radius, centred on each school's postcode centre point. This definition of boundaries is discretionary, pertaining to what in the context of the project was considered an adequate extent for the study area around each school. A boundary of this kind could be replaced by any other, without calling into question the adopted method. A number of morphological indicators were computed for the built environments within these areas, based on an open access vector dataset (OS, 2015) describing the full road network hierarchy and the footprints of all buildings.

The chosen morphological indicators cover three fundamental aspects of urban structure, which change significantly across urban areas and historical periods (Figure 6). These are: the geometry and topology of the street network (i.e., streets and junctions), the geometry and topology of urban blocks, and the geometry, density and grain of buildings. We note however, that a larger number of morphological attributes could be

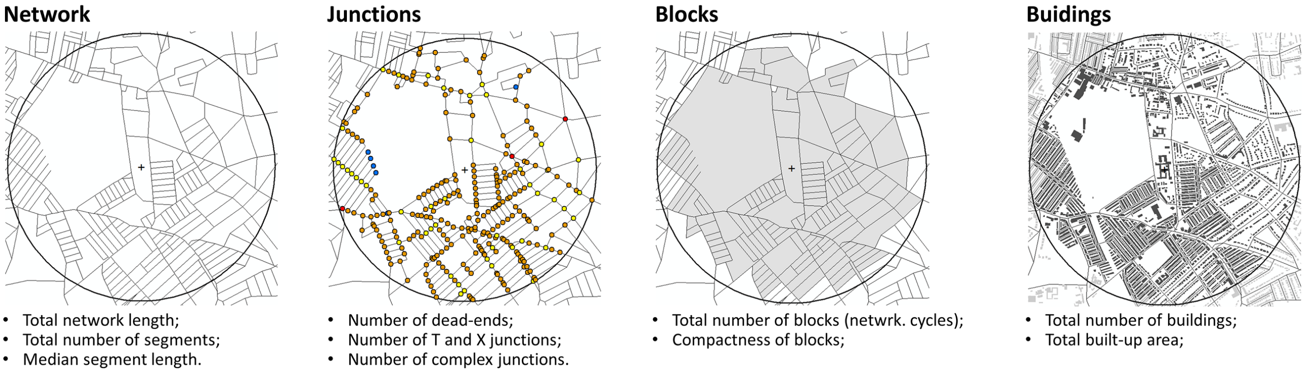


Figure 6. Graphic depiction of the adopted morphological indicators.

considered without jeopardizing the workings of the proposed method. The full list of the morphological indicators used in this study is as follows:

a) Network attributes: total network length (meters), total number of segments (straight street stretches) and ratio between the number of junctions and the number of segments.
b) Junction attributes: total number of junctions; number of T-junctions (three segments junction); number of X-junctions (four segments junction); number of complex junctions (more than four segments) and number of dead-ends.
c) Block attributes: total number of blocks (i.e., regions of space bounded by streets) and compactness of blocks (standardized area/perimeter ratio, yielding 1 for a circle and values close to 0 for thin, elongated shapes).
d) Building attributes: total number of buildings (i.e., count of buildings footprints polygons), total built-up area (i.e., sum of buildings footprints) and ratio between the built-up area and the number of buildings (contiguous buildings are represented as single polygons, being counted as a single feature; therefore, the denser and continuous a built tissue is, the greater the value of this ratio).

Each morphological indicator results in a single figure for each area surrounding a school. After standardization of all indicators as Z-scores, a first screening for potential collinearities reduced their number to just three, non-correlated variables, namely: the ratio between the number of junctions and street segments, or [JunctSegs]; the compactness of urban blocks, or [CycCompct]; and the ratio between the built-up area and the number of buildings, or [AreaNBuild]. Multicollinearity between variables used for unsupervised classification exercises should be avoided (Tan, Steinbach, & Kumar, 2005). Still, the three remaining variables are capable of describing the connectivity of the street network (through the density of junctions per street segment, [JunctSegs]), the general geometric shape of urban blocks (through their compactness, [CycCompct]) and the density and continuity of buildings (through the ratio [AreaNbuild], by the reasons mentioned before). These variables are then subjected to hierarchical cluster analysis.

Cluster analysis is a family of unsupervised classification methods aimed at dividing data into homogeneous classes (or clusters), so that the objects in a given class are more similar among themselves than to the objects in the other classes. It differs from supervised classification techniques (i.e., classification with a previous model or classification label), deriving the resulting classes only from the data itself; that is, as the result of intrinsic cleavages and associations between the data points and not of pre-defined classification criteria. Here we used Ward's minimum variance method (Tan et al., 2005), one of the most common hierarchical classification algorithms.

Ward's method starts with all objects separated and each object being a cluster; at each iteration, the two clusters the merging of which would lead to the minimum increase in total within-cluster variance are joined, becoming a single cluster. The process continues until all objects are merged into a single cluster. Figure 7 shows the resulting dendrogram; the length of the vertical lines represents the value of the inter-cluster dissimilarity between each cluster's two predecessors. Thus, one should look for a cutting level of the dendrogram where the vertical lines are all long and at which the number of clusters is parsimonious. Figure 7 also shows that, in our case, a division into four clusters seems optimal and this is indeed confirmed by the cubic clustering criterion (CCC; Sarle, 1983), whose value peaks at four clusters.

Having extracted these four clusters, we inspect their profiles on the three morphological variables (i.e., [JunctSegs], [CycCompct] and [AreaNBuild]), as well as the urban tissues to which they correspond. Figure 8 shows, for each cluster, an image of the case that is closer to the cluster's centroid (which may be considered its 'archetype') and also a chart depicting the values of each cluster's members on the three morphological variables (where the archetypal cases are represented by a thicker line). Visual inspection of the maps of each cluster's archetype reveals evident differences between the four urban tissues they describe. Also, the values of the members of each cluster on the three morphological variables

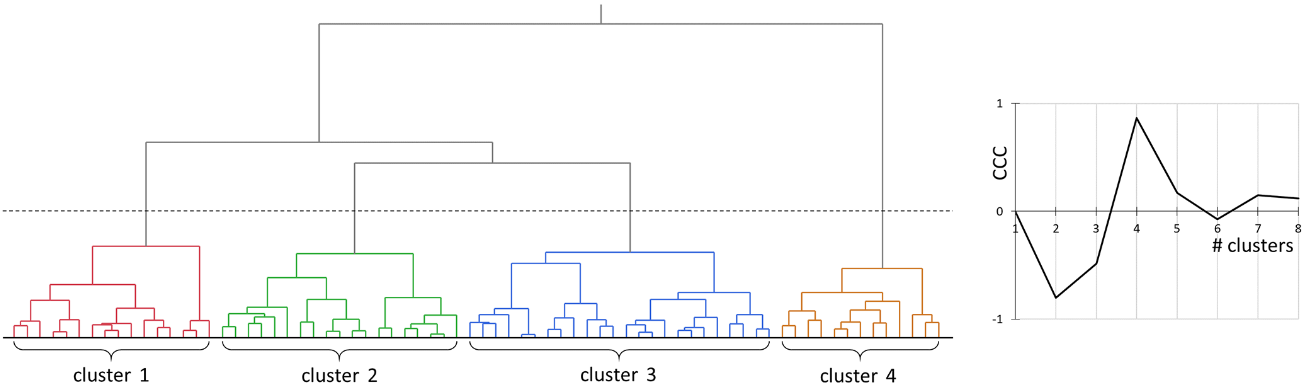


Figure 7. Hierarchical classification results.

(see lower charts on Figure 8) are quite similar within each cluster and clearly different between clusters. We can thus semantically characterize the four resulting morphological clusters in the following way:

Cluster 1 (n = 16): "Modern planned areas 1", composed of large and geometrically regular blocks (high average [CycCompct]). There are free-standing small buildings, creating a sparse urbanscape (low average [AreaNBuild]). There are very sparse street grids, with long street segments and few intersections (low average [JunctSegs]).

Cluster 2 (n = 20): "Modern planned areas 2", with blocks similar to cluster 1 (same average [CycCompct]). However, buildings (even if also small and separated) are more numerous in relation to Cluster 1 (higher [AreaNBuild]). The main difference between the two clusters is that in this one the street grid is denser, with more frequent junctions and shorter street segments (significantly higher average [JunctSegs]).

Cluster 3 (n = 25): "Early suburban developments", with more irregular and smaller urban blocks (lower [CycCompct]), and a more organic street grid. Buildings, although still mostly separated, are more densely organized (higher average [AreaNBuild], with greater variance). The average junction density ([JunctSegs]) is slightly lower than cluster 2, but its variance is higher.

Cluster 4 (n = 16): "Central historical fabric", with small elongated blocks (low [CycCompct]). There are densely packed, contiguous buildings (high [AreaNBuild]) in an organic and very dense street grid (high [JunctSegs]).

These morphological differences between the four clusters should correspond to also different epochs of urban expansion, with cluster 4 representing the older urban tissues and cluster 1 contemporary ones. This is indeed confirmed by looking at the spatial distribution of the four morphological clusters over the Liverpool City Region. Figure 9 shows that the members of cluster 4 are all located in the central area of the City of Liverpool and only on the east bank of the Mersey (which is the area of oldest occupation). Cluster 3 members immediately surround this central area, while appearing also on the west bank of the Mersey (i.e., on the Wirral Peninsula). On the right bank, Cluster 2 and 1 have mainly peripheral locations, with a greater incidence of cluster 1 on the farthest areas. On the Wirral peninsula this pattern is not so clear, because urbanization there is more recent and not so intensive.

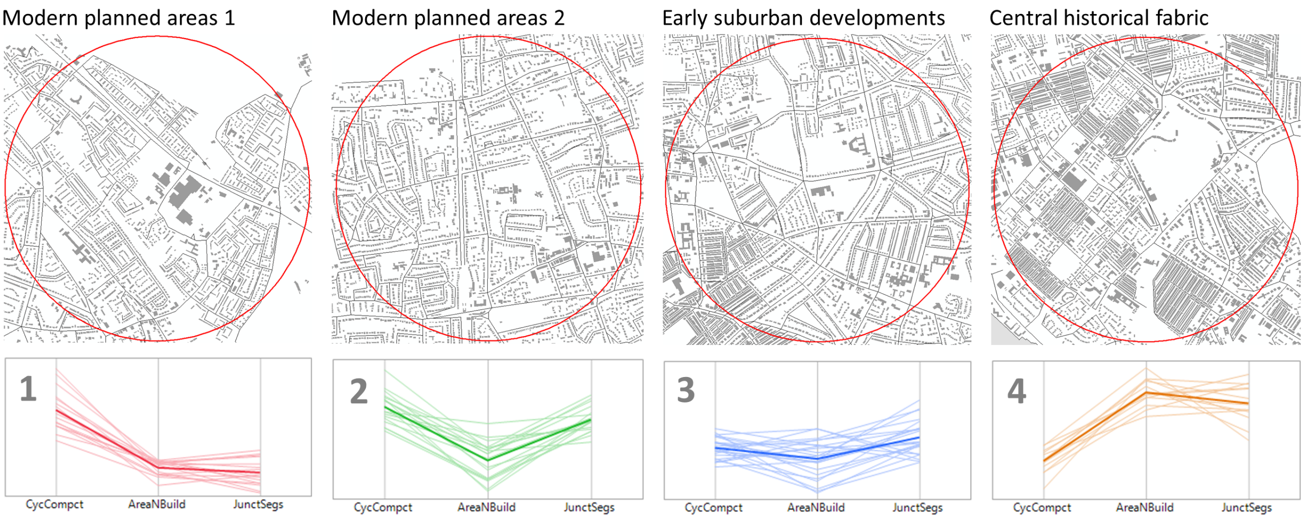


Figure 8. Archetypes and numerical profiles of the four morphological clusters.

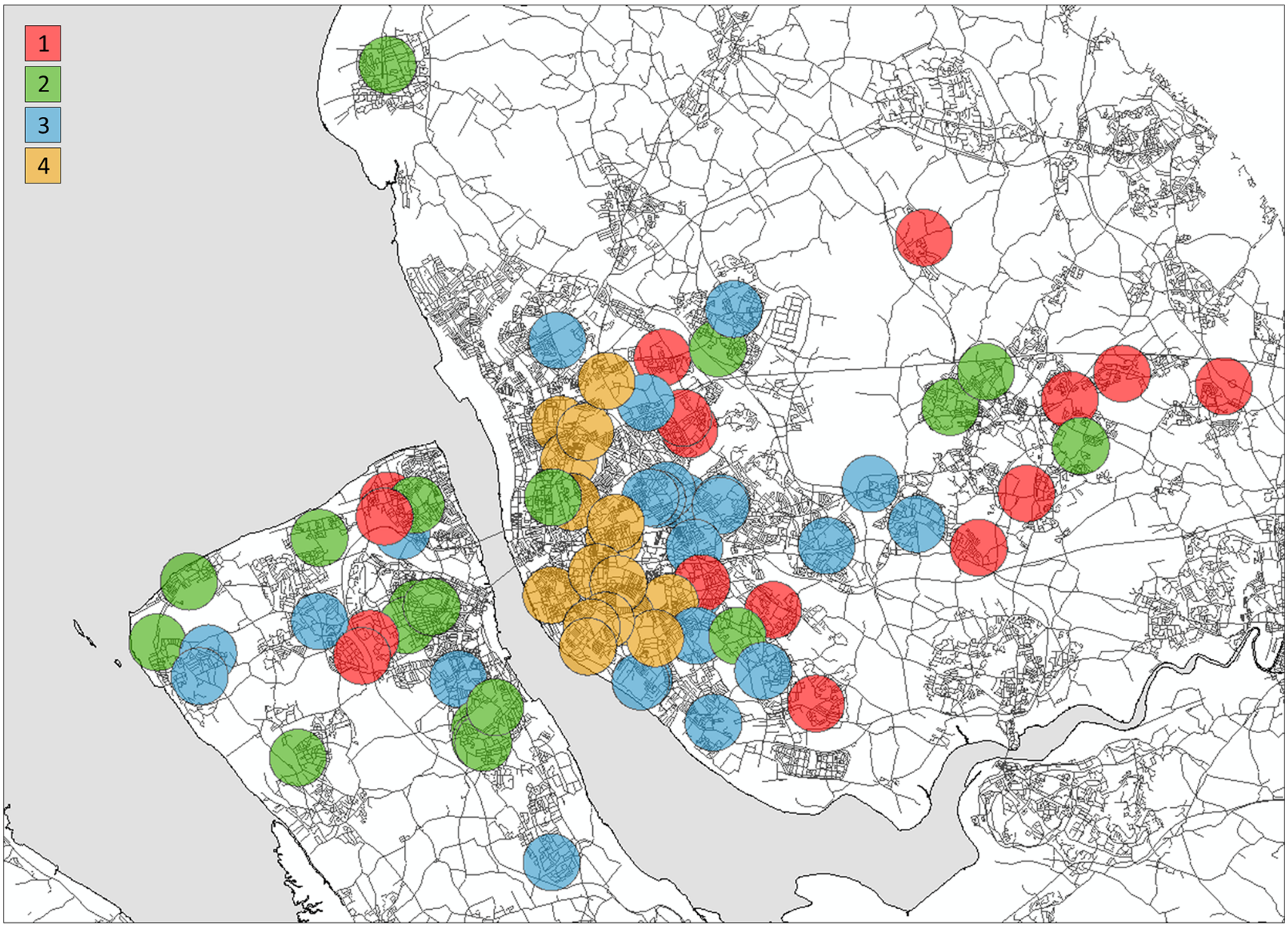


Figure 9. Geographical distribution of the four morphological clusters.

The proposed method for characterizing morphological urban contexts is therefore able to reduce the large initial variability of urban forms to a compact, yet meaningful categorization, of just four types of contexts. It takes into account detailed aspects of urban form, capable of detecting differences between urban tissues of different epochs and phases of development, which are otherwise difficult to classify objectively.

We summarize the morphological characterization of the 77 secondary schools by adding another layer to the socio-economic information displayed on Figure 5 (see Figure 10). For each school (i.e., each column of Figure 10) we record the morphological cluster to which it belongs, as well as its values (z-scores) on the three morphological variables that were used to define the clusters. With all the information produced by the proposed methods thus summarized, we can now use it to generate several context-informed purposive samples, aimed at different research questions and objectives.

5. Designing Context-Informed Purposive Samples of Urban Objects

In this section we use the data of the context characterization methods described before, in order to simulate four types of purposive samples proposed by Patton (1990), namely: *maximum variation sample*, *intensity sample* and two different types of *homogeneity samples* (Figure 11). Each of these illustrative samples is constituted by 16 observation units. This was the sample size that was used in the research project mentioned before, representing 23% of the whole population and being commensurate with the typical sample sizes of purposeful sampling.

When the population to be sampled is also small (as it is the case of the 77 Merseyside's secondary schools), random samples may not be an adequate way of achieving representativeness of the studied phenomena. In such cases a random sample has a non-negligible probability of not being representative at all, and a maximum variation sample may be a more efficient way of achieving representativeness. Maximum variation samples capture the extremes of a given set of characteristics, relevant for the problem under consideration. The logic behind this type of sampling is that any potential patterns found across the cases of such a sample, derive their significance from having emerged out of maximal heterogeneity (Patton, 1990).

In order to generate a maximum variation sample, we start by defining the dimensions along which varia-

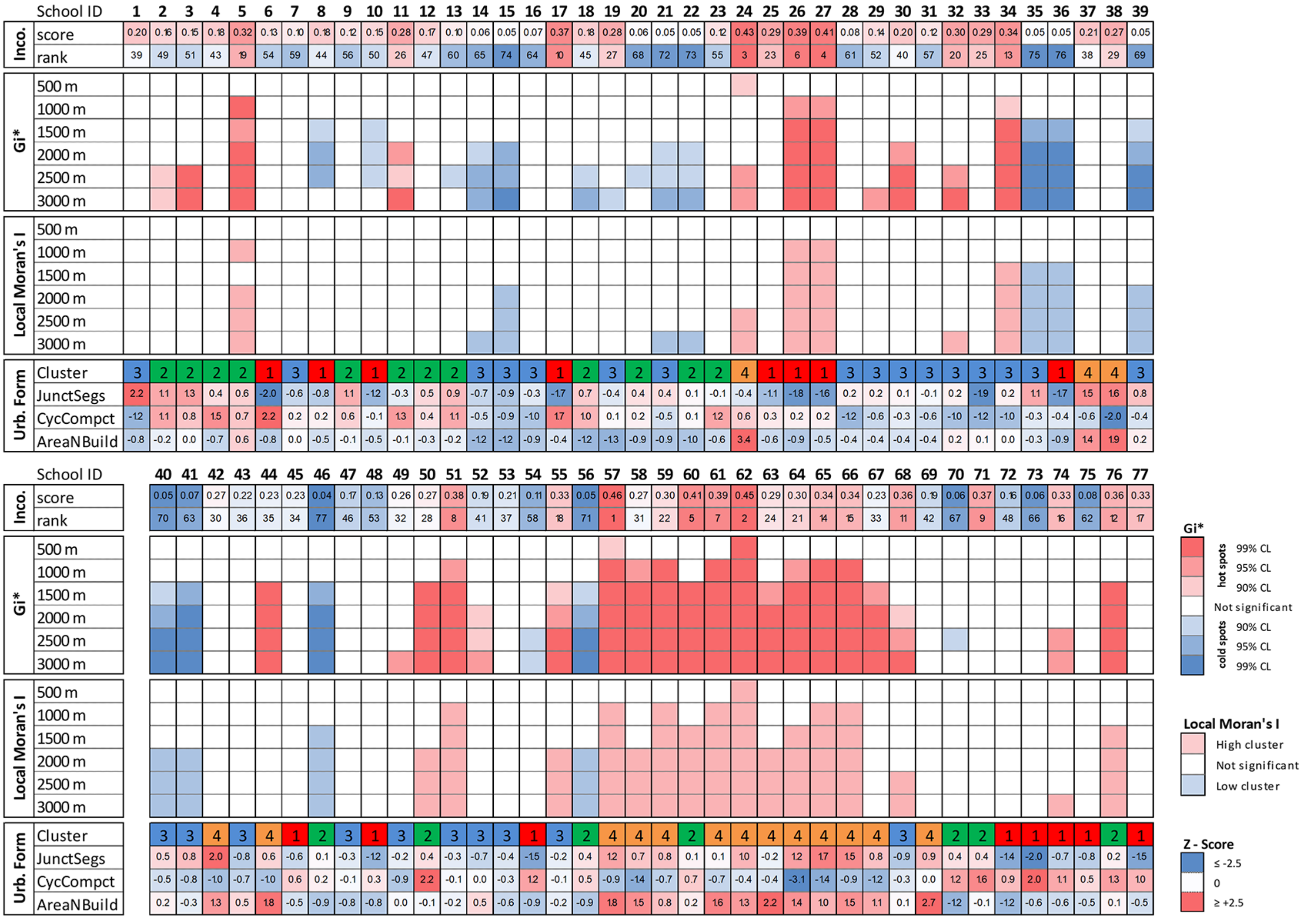


Figure 10. Social and physical contexts of the 77 secondary schools.

tion will be maximized; in our case, these are the social and physical composition of the urban contexts of the secondary schools. We have two extremes regarding social contexts: deprived and non-deprived, as defined in Section 3 (i.e., when such characteristics are verified across several spatial scales at once). Regarding material contexts, we have four possible types of variation, namely the four morphological clusters defined in Section 4. A maximum variation sample of urban contexts along these two dimensions with 16 observations, would therefore be composed by four cases of each morphological cluster, namely the two most deprived and the two least deprived (see Figure 11a).

Such a sample, covering the extremes of social and morphological variation observed in the population, would allow the investigation of the following research questions:

- Which regularities (if any) may be observed across all cases? These may be deemed general or transversal phenomena, independent of both social and morphological contexts.
- Which regularities (if any) may be observed across the four cases of each morphological cluster, independently of their social composition? These may be attributable to specific physical characteristics.
- Which regularities (if any) may be observed only on deprived and/or affluent cases, independently of their specific morphology? These may be attributable to specific social characteristics.

Intensity sampling aims at selecting cases in which the intensity of the phenomenon under investigation is maximized (Patton, 1990). In contrast to maximum variation, intensity sampling presupposes a previous observation or hypothesis to be further explored. For example, looking at Figure 10, it is clear that the majority of cases in cluster 4 are highly income-deprived (11 out of 16 cases), whereas the cases in cluster 1 tend to be rather less deprived (only 3 out of 16 cases showing high income deprivation). Independently of the causes behind such regularity, one may argue that high income deprivation is typical of the morphological contexts described by cluster 4 and atypical of those described by cluster 1. Furthermore, these two clusters are clearly the most separated in time and most dissimilar in morphological terms, with cluster 4 representing historical urban tissues and cluster 1 representing modern planned ones of the urban sprawl type. Thus, a sample composed of cluster's 4 deprived cases and cluster's 1 non-deprived cases would maximize the intensity of both social and morphological differences.

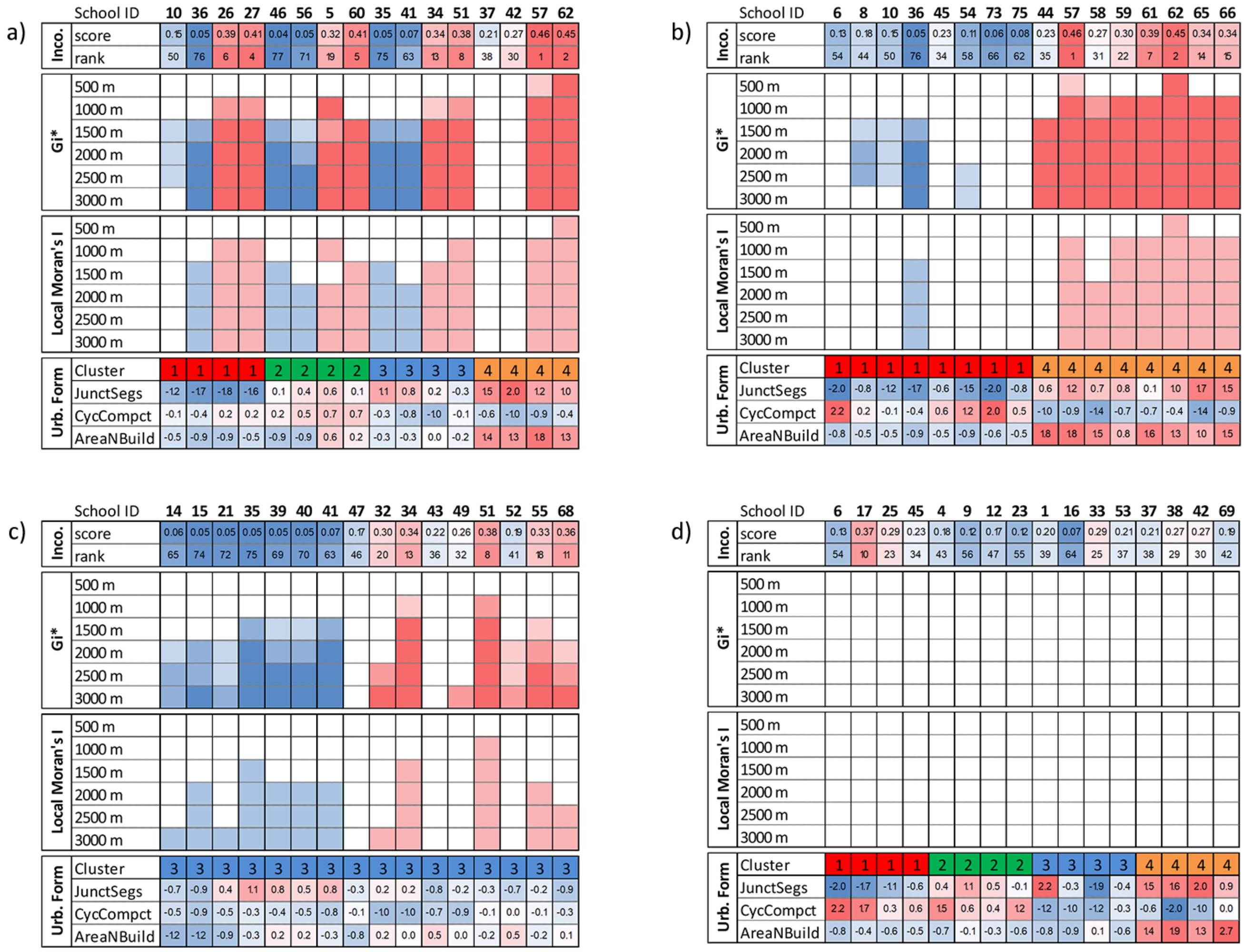


Figure 11. Four types of purposive samples: a) maximum variation sample, b) intensity, c) morphological homogeneity and d) socio-economic homogeneity samples.

An intensity sample with 16 observations would therefore be composed by the 8 most deprived cases of cluster 4, and the 8 least deprived cases of cluster 1 (see Figure 11, b). Such a sample would maximize both the probability of observing urban community inequalities and the intensity of their specific characteristics. It would allow the investigation of the following research questions:

- Which factors may explain the observed association between the two types of physical contexts and their specific social compositions?
- Which regularities (if any) are specific to each group of deprived and non-deprived cases?
- To which extent the specific material contexts of deprived and non-deprived cases are related to such regularities?

Finally, we propose two variations of 'homogeneity sampling' (Patton, 1990), a strategy which is the opposite of maximum variation sampling. Instead of maximizing variation, one tries to minimize it on one or several variables of interest. The purpose here is to study a given subgroup in depth, or to maintain the variability of a given dimension constant in order to reduce as far as possible its potential confounding effects.

We generate two different homogeneous samples: one in which we maintain the physical characteristics constant (Figure 11c); and another one in which we do the same regarding socio-economic characteristics (Figure 11d). In the first case, we select only cases belonging to cluster 3, half of them highly deprived and another half affluent. We choose to hold cluster 3 constant, because it is the most frequent morphological type (n = 25) with a high socio-economic variability, which we try to maximize by selecting only highly deprived and affluent cases. The objective is to study specifically the impacts of socio-economic composition of urban contexts, while maintaining urban form constant. Because all cases have similar physical contexts, we can be reasonably confident that any detected regularities would pertain to the socio-economic characteristics of the selected cases. Such a sample would allow the investigation of the following research question:

- Which are the specific impacts of deprived and affluent socio-economic contexts on urban communities, while controlling for urban form's potential confounding effects?

In the second case (Figure 11d), we select only cases showing not-significant spatial clustering of either deprivation or affluence, with income scores close to the mean, while choosing four cases of each morphological cluster. Conversely to the previous situation, the objective here is to study the specific impacts of physical characteristics, while maintaining socio-economic contexts constant and at an average level (i.e., neither particularly deprived nor affluent). Again, because all cases have similarly average socio-economic characteristics, but also quite different morphological contexts, one would expect that any regularities found would pertain to differences in physical context. This sample would allow the investigation of the following research question:

- Which are the specific impacts of different morphological contexts on urban communities, while controlling for potential socio-economic confounding effects?

We end this section by displaying the spatial distributions of the four sample simulations discussed above (Figure 12). Different samples result in also different spatial distributions, covering diverse parts of Liverpool City Region. Each sample serves different research objectives and none is a priori preferable over the others.

6. Conclusions

This article has proposed a set of GIS methods for quantifying, classifying and sampling the social and physical urban contexts of 77 secondary schools in Liverpool City Region, Merseyside, UK. The proposed methods overcome a number of shortcomings that current approaches to the characterization of urban contexts suffer from, namely: the exposure to the MAUP and its biasing effects; the rudimentary level at which urban form is commonly quantified and classified; and the lack of methodology in supporting purposive sampling, for exploring the complex relationships between urban contextual characteristics and other variables of interest.

Regarding the characterization of social urban contexts, and as the means to overcome the deleterious effects of MAUP, we make use of LISA, applied to avail-

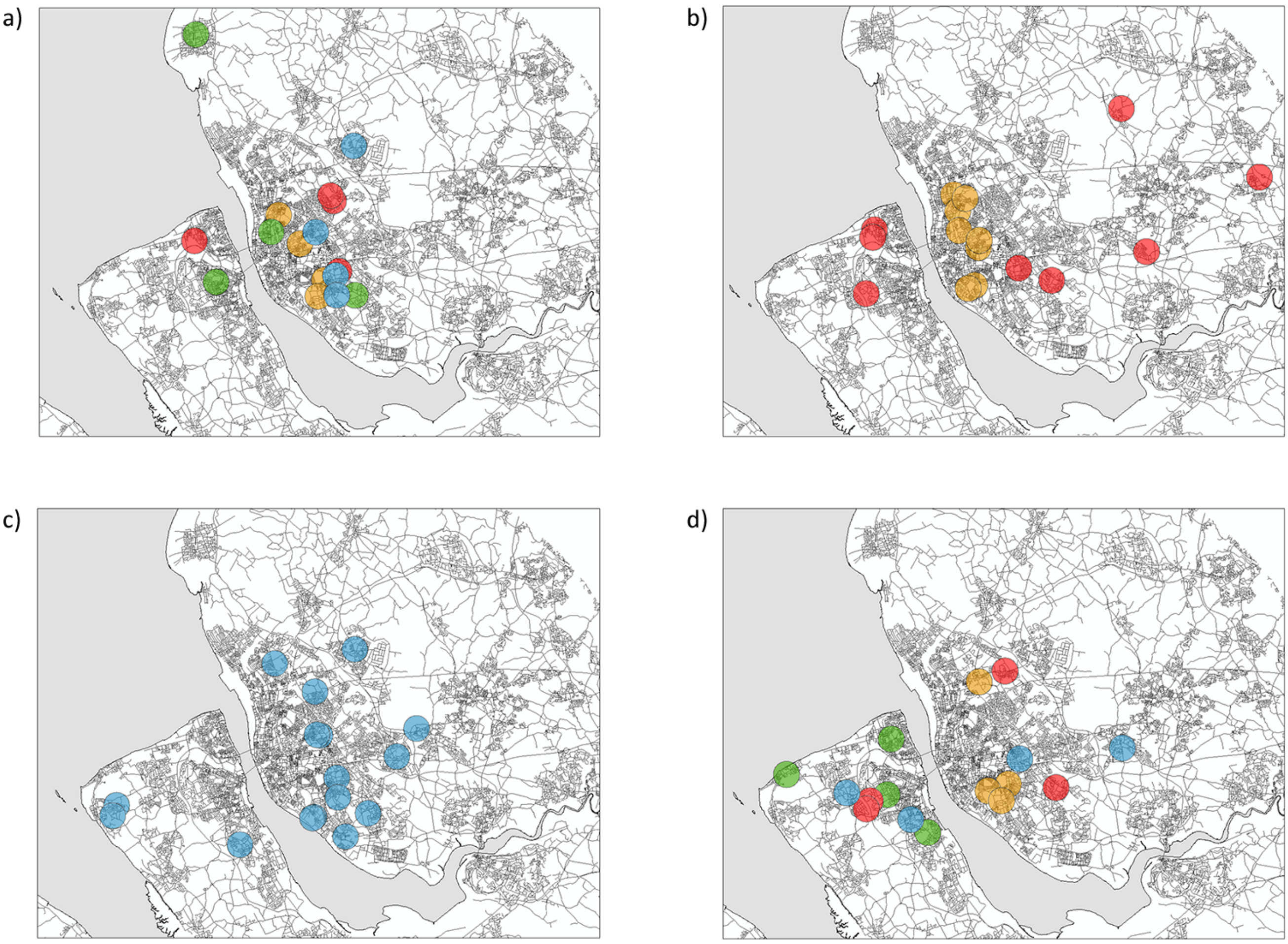


Figure 12. Geographical distributions of the four sample types; a) maximum variation, b) intensity, c) morphological homogeneity and d) socio-economic homogeneity.

able socio-economic indicators. We draw attention to the importance of evaluating the consistency of socio-economic indicators across several spatial scales, in order to identify accurately local areas where such indicators attain consistent high or low scores, as well as others where they do not. We apply simultaneously two LISA methods, namely the Gi* statistic and the Local Moran's I statistic, showing how their conjoint use is capable of providing detailed information about the specific social context of each school.

The physical characteristics of urban contexts are quantified through three morphological variables measured in GIS, namely the ratio between the number of junctions and street segments, the general geometric shape of urban blocks, and the density and continuity of buildings. We then use cluster analysis to objectively classify the physical context of each school, into a compact, yet meaningful categorization, of just four types of contexts: "modern planned areas 1", "modern planned areas 2", "early suburban developments" and "central historical fabric", each corresponding to different periods of urban expansion and types of geographical distribution in the history of Liverpool. By dividing data into classes that are derived by algorithmic means from the data themselves, this method overcomes the potential bias of pre-existing semantic classifications, while resulting in a high level of morphological detail.

Finally, the data generated by these methods is summarized into visualization schemes, revealing the relative variation of the social and physical contexts of the 77 schools. We use such schemes to produce four types of purposive samples, illustrating the design of context-informed samples of urban objects, aimed at different potential research questions in community and neighbourhood studies. We note that purposive sampling strategies, even though generally overlooked, can be extremely useful for exploring the inherently complex relationships between urban context and other variables of interest. The current focus on probabilistic sampling techniques, in its endeavour to find generalizable effects, is perhaps not the best initial approach to such intricate and elusive phenomena. We suggest that purposive sampling strategies, by virtue of selecting specific information-rich cases, may be more fruitful for exploring the potential impacts of different urban contexts, whose generality may subsequently be tested with larger probabilistic samples.

This work responded to the research objectives of visualising and measuring social inequalities in Liverpool's urban environments as part of a specific research project. However, the proposed methods are not limited to the chosen variables or urban objects (schools), do not depend on geographical context and can address a larger range of dimensions without loss of consistency. On the contrary, they can provide a robust and efficient methodology on comparative profiling and sampling of a wide range of socio-economic factors and urban forms, across time, scale and contexts.

Acknowledgements

The research project "Visualizing Inequalities in Community Networks to Enhance Participatory Planning" was supported by the Leverhulme Trust Research Project Grant RPG-2014-169. Miguel Serra is supported by the FCT grant SFRH/BPD/111260/2015.

Conflict of Interests

The authors declare no conflict of interests.

References

Anselin, L. (1995). Local Indicators of Spatial Association—LISA. *Geographical Analysis*, *27*(2), 93–115.

Barthelemy, M. (2015). From paths to blocks: New measures for street patterns. *Environment and Planning B: Urban Analytics and City Science*, *44*(2), 256–271.

Caughy, M., Leonard, T., Beron, K., & Murdoch, J. (2013). Defining neighborhood boundaries in studies of spatial dependence in child behaviour problems. *International Journal of Health Geographics*, *11*(24), 1–12.

Charron, M. (2009). Neighbourhood characteristics and the distribution of police-reported crime in the city of Toronto. *Crime and Justice Research Paper Series*. Ottawa: Statistics Canada.

Crane, R. (2000). The influence of urban form on travel: An interpretive review. *Journal of Planning Literature*, *15*(1), 3–23.

Cummins, S., Macintyre, S., Davidson, S., & Ellaway, A. (2005). Measuring neighbourhood social and material context: Generation and interpretation of ecological data from routine and non-routine sources. *Health & Place*, *11*, 249–260.

Cutrona, C., Wallace, G., & Wesner, K. (2006). Neighborhood characteristics and depression: An examination of stress processes. *Current Directions in Psychological Science*, *15*(4), 188–192.

DCLG (2015). *The English indices of deprivation 2015—Technical report*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/464485/English_Indices_of_Deprivation_2015_-_Technical-Report.pdf

Dibble, J., Prelorendjos, A., Romice, O., Zanella, M., Strano, E., Pagel, M., & Porta, S. (2017). On the origin of spaces: Morphometric foundations of urban form evolution. *Environment and Planning B: Urban Analytics and City Science*. https://doi.org/10.1177/2399808317725075

ESRI. (2011). ArcGIS Desktop: Release 10.3: Redlands, CA: Environmental Systems Research Institute.

Gambaro, L., Joshi, H., Lupton, R., Fenton, A., & Lennon, M. C. (2016). Developing better measures of neighbourhood characteristics and change for use in studies of residential mobility: A case study of Britain in the early 2000s. *Applied Spatial Analysis*, *9*, 569–590.

Gil, J., Beirao, J., Montenegro, N., & Duarte, J. (2011). On the discovery of urban typologies: Data mining the many dimensions of urban form. *Urban Morphology*, *16*(1), 27–40.

Hamaina, R., Leduc, T., & Moreau, G. (2012). Towards urban fabrics characterization based on buildings footprints. In J. Gensel, D. Josselin, & D. Vandenbroucke (Eds.), *Bridging the Geographic Information Sciences, Lecture Notes in Geoinformation and Cartography (LNG&C)* (pp. 327–346). Berlin and Heidelberg: Springer.

Inoue, Y., Stickley, A., Yazawa, A., & Shirai, K. (2016). Neighbourhood Characteristics and Cardiovascular Risk among Older People in Japan: Findings from the JAGES Project. *PloSONE*, *11*(10), 1–16.

Kim, R., Ali, M., Sur, D., Khatib, A., & Wierzba, T. (2012). Determining optimal neighborhood size for ecological studies using leaving-one-out cross validation. *International Journal of Health Geographics*, *11*(10), 1–6.

LCC. (2015). *The index of multiple deprivation 2015: A Liverpool analysis*. Liverpool: Liverpool City Council.

Lebel, A., Pampalon, R., & Villeneuve, P. (2007). A multiperspective approach for defining neighbourhood units in the context of a study on health inequalities in the Quebec City region. *International Journal of Health Geographics*, *6*(27), 1–15.

Lupton, R. (2003). *Neighbourhood effects: Can we measure them and does it matter?* (CASE paper 73). London: Centre for Analysis of Social Exclusion, London School of Economics.

MacDonald, J., Wise, M., & Harris, P. (2008). *The health impacts of the urban form: A review of reviews*. Retrieved from http://hiaconnect.edu.au/old/files/Health_Impact_of_Urban_Form.pdf

Marshall, S. (2005). *Streets & patterns*. London and New York: Spon Press.

Miles, R., Coutts, C., & Mohamadi, A. (2011). Neighborhood urban form, social environment and depression. *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, *89*(1), 1–18.

O'Brien, J., Serra, M., Hudson-Smith, A., Psarra, S., Hunter, A., & Zaltz Austwick, M. (2016). Ensuring VGI credibility in urban-community data generation: A methodological research design. *Urban Planning*, *1*(2), 88–100.

O'Brien, J., Garcia Vélez, L., & Zaltz Austwick, M. (2017). Visualizing the impacts of movement infrastructures on social inclusion: Graph-based methods for observing community formations in contrasting geographic contexts. *Social Inclusion*, *5*(4), 132–146.

Office of National Statistics. (2011). Census geography: An overview of the various geographies used in the production of statistics collected via the UK census. Retrieved from https://www.ons.gov.uk/methodology/geography/ukgeographies/censusgeography

Openshaw, S. (1983). *The modifiable areal unit problem*. Norwick: Geo Books.

Ord, J., & Getis, A. (1995). Local spatial autocorrelation statistics: Distributional issues and an application. *Geographical Analysis*, *27*(4), 286–306.

OS. (2015). VectorMap Local. Dataset. Retrieved from https://www.ordnancesurvey.co.uk/business-and-government/products/os-open-map-local.html

Patton, M. (1990). *Qualitative evaluation and research methods*. Beverly Hills, CA: Sage.

Pont, M. B., & Haupt, P. (2010). *Spacematrix: Space, density and urban form*. Roterdam: NAi Publishers.

Rae, A. (2012). Spatially concentrated deprivation in England: An empirical assessment. *Regional Studies*, *46*(9), 1183–1199.

Sarle, W. (1983). *Cubic clustering criterion* (SAS Technical Report, 108). Cary, NC: Statistical Analysis System Institute.

Serra, M., Gil, J., & Pinho, P. (2016). Towards an understanding of morphogenesis in metropolitan street-networks. *Environment and Planning B: Urban Analytics and City Science*, *44*(2), 272–293.

Tan, P.-N., Steinbach, M., & Kumar, V. (2005). *Introduction to data mining* (1st ed.). Boston, MA: Addison-Wesley Longman Publishing.

Timperio, A., Jeffery, R., Crawford, D., Roberts, R., Giles-Conti, B., & Ball, K. (2010). Neighbourhood physical activity environments and adiposity in children and mothers: A three-year longitudinal study. *International Journal of Behavioral Nutrition and Physical Activity*, *7*(18), 1–8.

Townshend, T., & Lake, A. (2009). Obesogenic urban form: Theory, policy and practice. *Health & Place*, *15*, 909–916.

van Ham, M., & Manley, D. (2012). Neighbourhood effects research at a crossroads: Ten challenges for future research. *Environment and Planning A*, *44*, 2787–2793.

Vernez-Moudon, A. (1994). Getting to know the built landscape: Typomorphology. In A. Frank (Ed.), *Ordering space: Types in architecture and design* (pp. 289–314). New York: Van Nostrand Reinhold.

About the Authors

Miguel Serra is Senior Researcher at CITTA (Research Centre for Territory, Transports and Environment, University of Porto) and Honorary Research Associate at the Bartlett School of Architecture (University College London, UCL). His current research focuses on large spatial networks, quantitative urban morphological descriptions and classifications, geographic information systems and urban data analysis and visualization. He was Senior Research Associate at UCL (2015–2016) in the research project "Visualizing Inequality in Community Networks to Enhance Participatory Planning", funded by the Leverhulme Trust.

Sophia Psarra is Reader at UCL. Her latest book, *The Venice Variations* (UCL Press, 2018), explores spatial and social networks as multi-authored processes of formation, alongside the ways in which they interact with urban design based on individual intention. Psarra was co-investigator in the research project “Visualising Inequality in Community Networks to Enhance Participatory Planning” (UCL) funded by the Leverhulme Trust; and co-principal investigator in a project exploring spatial, social networks and organisational innovation, funded by the NSF (the University of Michigan).

Jamie O’Brien is Hon. Research Associate at Bartlett Centre for Advanced Spatial Analysis (CASA), University College London, and is a consultant in GIS data visualization. He was Senior Research Associate at CASA (2014–2017), where he focused on communities’ uses of local spatial assets in achieving their sense of place. He holds a Doctorate in Engineering from UCL’s Bartlett School of Architecture (supported by an EPSRC studentship; 2002–2006), and is Fellow of the Higher Education Academy.

Urban Planning (ISSN: 2183–7635)
2018, Volume 3, Issue 1, Pages 75–99
DOI: 10.17645/up.v3i1.1287

Article

#London2012: Towards Citizen-Contributed Urban Planning Through Sentiment Analysis of Twitter Data

Anna Kovacs-Gyori [1,*], Alina Ristea [1], Clemens Havas [1], Bernd Resch [1,2] and Pablo Cabrera-Barona [3,4]

[1] Department of Geoinformatics—Z_GIS, University of Salzburg, 5020 Salzburg, Austria; E-Mails: anna.gyori@sbg.ac.at (A.K.-G.), mihaela.ristea@sbg.ac.at (A.R.), clemensrudolf.havas@sbg.ac.at (C.H.), bernd.resch@sbg.ac.at (B.R.)
[2] Center for Geographic Analysis, Harvard University, Cambridge, MA 02138, USA; E-Mail: bresch@fas.harvard.edu
[3] Institute of Higher National Studies—IAEN, 170507 Quito, Ecuador; E-Mail: pablo.cabrera@iaen.edu.ec
[4] Latin American Social Sciences Institute—FLACSO Ecuador, 170201 Quito, Ecuador; E-Mail: pablo.cb@live.com

* Corresponding author

Submitted: 26 November 2017 | Accepted: 26 February 2018 | Published: 29 March 2018

Abstract
The dynamic nature of cities, understood as complex systems with a variety of concurring factors, poses significant challenges to urban analysis for supporting planning processes. This particularly applies to large urban events because their characteristics often contradict daily planning routines. Due to the availability of large amounts of data, social media offer the possibility for fine-scale spatial and temporal analysis in this context, especially regarding public emotions related to varied topics. Thus, this article proposes a combined approach for analyzing large sports events considering event days vs comparison days (before or after the event) and different user groups (residents vs visitors), as well as integrating sentiment analysis and topic extraction. Our results based on various analyses of tweets demonstrate that different spatial and temporal patterns can be identified, clearly distinguishing both residents and visitors, along with positive or negative sentiment. Furthermore, we could assign tweets to specific urban events or extract topics related to the transportation infrastructure. Although the results are potentially able to support urban planning processes of large events, the approach still shows some limitations including well-known biases in social media or shortcomings in identifying the user groups and in the topic modeling approach.

Keywords
geographic information; GIS; Olympic Games; planned events; sentiment analysis; social media; spatiotemporal analysis; topic extraction

Issue
This article is part of the issue "Crowdsourced Data and Social Media for Participatory Urban Planning", edited by Bernd Resch (University of Salzburg, Austria), Peter Zeile (Karlsruhe Institute of Technology, Germany) and Ourania Kounadi (University of Salzburg, Austria).

1. Introduction

Cities are complex systems (Castells, 1996; Hall, 1966; Theodore, 2006), consisting of two main elements: the people as residents or visitors, and the infrastructure to fulfill their needs ranging from housing to recreation or even self-realization (Costanza et al., 2007; Maslow, 1943). Some of the infrastructure or related networks are static and mostly physical, such as buildings or the road and electricity networks, whereas others are more dynamic, like social, transportation, or financial networks. From an urban analysis viewpoint, the dynamic nature of these systems is challenging, especially in the case of large cities with millions of people constantly on the move and having different needs and preferences. These challenges do not only result from the sheer amount of people, but also from the intense spatiotemporal variability originating from urban dynamism and from the constantly changing subjective needs of each person. Therefore, effective planning practice requires analysis at high spatial and temporal scales to understand this dynamism of urban life and processes.

Traditional methods, such as questionnaires or counting, are not capable of handling such fine temporal and spatial scales at all, or they are highly resource-consuming and, therefore, slow and costly, and thus not up-to-date. This is where the advantages of the data-driven era become relevant, most concretely with respect to the real-time availability of social media data. This data provides unseen contextual insights into spatiotemporal phenomena on a finer scale in cities through users' digital traces on different online platforms such as Twitter, Foursquare/Swarm and Flickr (Abbasi, Rashidi, Maghrebi, & Waller, 2015; Aubrecht, Ungar, & Freire, 2011; Crooks et al., 2015; Girardin, Vaccari, Gerber, Biderman, & Ratti, 2009). This is of central importance to urban planning dealing with the optimization of the above-mentioned networks.

Planners are responsible for land use strategies, the design of public places, or transportation planning, which constitute essential factors of urban life (McGill, 2017). An aspect of particular importance for urban planning is the investigation of the effects of a planned large event, considering residents and visitors. These events have a special role in planning because they are usually temporary and require completely different circumstances and conditions compared to the average daily routines of urban life. In contrast with other unplanned events such as emergencies (e.g., natural, industrial and manmade disasters), direct preparations can be made for planned events, not just precautionary measures (Getz & Page, 2016). As a consequence, they are preceded by extensive planning and preparation efforts; but such events frequently still face severe inconveniences or even disruptions, most strikingly with respect to the transportation of people, presumably in different ways for residents and visitors. Therefore, the distinction between these two groups is crucial in most of the analyses due to their different needs, behavioral patterns, and exposure to the effect of a large planned event.

Thus, analyses at fine temporal scales are inevitable for examining citizens' mobility, which can help in detecting patterns and anomalies through the understanding of underlying problems or phenomena in the context of urban transportation. For instance, citizens' trajectories and the number of people moving through the city vary over time during the day, but also between days depending on the weather, weekday, planned and sudden events, traffic density, and many other factors (Sagl, Resch, Hawelka, & Beinat, 2012). Therefore, traditional annual commuting statistics are not informative on such fine spatial and temporal scales because they are mostly produced only once or twice per year and aggregated to spatial planning units. First, this results in commuting data that do not actually reflect real travel directions due to their aggregated nature, and, second, everyday individual trajectory details are lost through the aggregation.

As a consequence, social media, providing digital spatiotemporal traces of individuals, grant valuable mobility information, particularly through their nature of a large and continuous source of data and their fine scale in space and time. Another advantage of using these sources is the potential for extracting direct feedback about city life-related topics, places or phenomena by revealing subjective aspects as well, such as public mood or emotions (Frank, Mitchell, Dodds, & Danforth, 2013; Quercia, Ellis, Capra, & Crowcroft, 2012; Resch, Summa, Zeile, & Strube, 2016). This provides an opportunity to investigate what people actually think about parts of the city and the direct or indirect effects of a large event. Thus, several methods and analyses have been developed in the area of opinion mining (e.g., Pak & Paroubek, 2010) and semantic topic extraction (e.g., Steiger, Westerholt, Resch, & Zipf, 2015) for use in urban planning.

However, to the best of our knowledge, limited research has been conducted to analyze social media data regarding planned large events, considering comparison days (before or after the event), different user groups (residents vs visitors) along with the linkage between sentiment analysis and topic extraction in one study. In our work, we intend to integrate all of these aspects to provide valuable knowledge about urban events, whereby our case study focuses on the 2012 Olympic Games in London. By exploring emotions and events in a city through social media analysis, we aspire to a better understanding of citizens' behaviors and needs in cities. Thereby, we can provide a basis to aid planners in identifying more specific urban planning issues for further in-depth analysis. In line with these goals, we intend to answer the following research questions in this article:

RQ1 → How can we identify distinctive characteristics of tweeting behavior in terms of spatiotemporal patterns and sentiments between "residents" and "visitors"?

RQ2 → Are there detectable changes in the spatial and temporal patterns, and sentiment of the tweets during the London Olympic Games compared to the days before and after it?

RQ3 → Which topics that are related to urban planning in the context of a large sports event can be identified through semantic analysis of social media posts?

2. Related Work

2.1. Citizen-Contributed Geographic Information to Describe Urban (and Spatial) Practices

Among the practical applications of geographic data extracted from social media, we can distinguish two main categories: quantitative and qualitative aspects.

Quantitative approaches describe spatiotemporal phenomena by leveraging the advantage of fine spatial and temporal scales of the data, such as crowdsourcing urban form and function (Crooks et al., 2015), or characterizing and classifying urban areas and location types (Noulas, Scellato, Mascolo, & Pontil, 2011).

Among quantitative analyses, mobility forms its own well-defined category. The applications range from the description of general spatiotemporal dynamics to revealing anomalies on urban (Aubrecht et al., 2011; Fujisaka, Lee, & Sumiya, 2010; Hasan & Ukkusuri, 2014), or global scales (Hawelka et al., 2014). Traffic and mobility modeling purposes are also present, such as validating travel demand models (Lee, Gao, & Goulias, 2016), and analyzing origin-destination flows (Cebelak, 2013).

There are also a few applications where researchers assess urban life from a *qualitative* point of view using different social media sources. Girardin et al. (2009) evaluated urban attractiveness by analyzing images from Flickr and mobile phone usage data, while Sun, Fan, Bakillah and Zipf (2015) used geo-tagged images for road-based travel recommendations. As another approach, several researchers improved and refined various methodologies for extracting emotions (Resch et al., 2016), transit rider satisfaction (Collins, Hasan, & Ukkusuri, 2013), and community happiness (Quercia et al., 2012) from Twitter data, also combined with demographics and other objective characteristics of a place such as education or obesity (Mitchell, Frank, Harris, Dodds, & Danforth, 2013), or even defined sentiment as a function of movement (Frank et al., 2013). The advantages of utilizing available additional datasets such as demographics, mobile phone data or mobility trajectories are twofold; they can help the interpretation of the primary results extracted from social media, and, on the other hand, they are also appropriate for validation purposes.

Further related approaches vary in their purpose, i.e., they are not specifically tailored to urban planning, but they can still be used in this context. They include classifying Twitter users (Pennacchiotti & Popescu, 2011), or describing geotemporal demographics (Longley, Adnan, & Lansley, 2015). Furthermore, the analysis of textual content to extract spatial information is becoming increasingly important (Cheng, Caverlee, & Lee, 2010; Dalvi, Kumar, & Pang, 2012; Kinsella, Murdock, & Hare, 2011; Li, Serdyukov, de Vries, Eickhoff, & Larson, 2011), including the generation of ambient geographic information (Stefanidis, Crooks, & Radzikowski, 2013), or the definition of geotag gazetteers (Keßler, Maué, Heuer, & Bartoschek, 2009).

2.2. Urban Planning, Social Media and Planned Large Events

Previous studies have shown that social media usage is generally more intensive during large events, and a concentration around the venue and impact on transportation is also identifiable (Gupta & Kumaraguru, 2012; Zhang, Ni, He, & Gao, 2016). The Olympic Games are considered one of the world's largest events, involving a lot of organizational tasks from social, technical, environmental, economic, demographic and transportation-related perspectives (Chen, 2012; Cook & Ward, 2011; Malfas, Houlihan, & Theodoraki, 2004).

Furthermore, a growing body of literature is dealing with the use of Twitter data for investigating the characteristics of different types of events, which is of importance also to the field of urban planning. Recently, a number of machine-learning approaches have been used to investigate electoral predictions (Gayo-Avello, 2013), stock market flows (Zhang, Fuehres, & Gloor, 2011), flu trends (Culotta, 2010; Ritterman, Osborne, & Klein, 2009), natural disasters (Fraustino, Liu, & Jin, 2012; Resch, Usländer, & Havas, 2017), or to detect large events (Lee & Sumiya, 2010; Li, Lei, Khadiwala, & Chang, 2012; Weng & Lee, 2011), even in near real-time (Zhao, Zhong, Wickramasuriya, & Vasudevan, 2011) and with respect to their impacts (Panteras et al., 2015).

Large sports events like the FIFA World Cup can also be identified using the content of the tweets, hashtags and distribution of retweets. Kim et al. (2015) applied topic modeling before and during the event, while Corney, Martin and Göker (2014) identified phrases (word n-grams) that showed a sudden increase in frequency in the dataset and then selected co-occurring n-grams to identify topics. By using sentiment analysis, researchers identified relationships between the public mood and large socioeconomic events in the media (Bollen, Mao, & Pepe, 2011), together with trends and possible predictions of the disposition theory (Yu & Wang, 2015), such as fanship for sports. Clearly, these analyses identified changes in activity patterns (e.g., supporters induce a general increase in number of tweets), in topic diversity, and in the spatial distribution of topics related to the event.

3. Data

The study area for the present analysis is Greater London, which has an expansion of 3,458 km^2. The Twitter data was obtained using the Twitter Streaming Application Programming Interface (Twitter INC, 2017) for the year 2012, and consists of tweet content and attributes such as user name, user location, and message time. We only harvested geolocated tweets, as our study requires geospatial and temporal analysis. To the best of our knowledge, this database does not contain retweets. It shall be noted that due to user practice and the policy of Twitter, in general, the tweets containing coordinates represent only a smaller subset of all tweets posted in a given period, about 1–10% according to previous studies (Morstatter, Pfeffer, Liu, & Carley, 2013; Zhang et al., 2016). Moreover, they are not evenly spread in space and among user groups, as youngsters tend to use social media more actively (Li, Goodchild, & Xu, 2013; Resch et al., 2017). These issues have been thoroughly discussed in existing literature (Steiger et al., 2015; Sui & Goodchild, 2011) and will be further detailed in the Discussion section. In addition, we want to point out that distinguishing personal and non-personal Twitter accounts was beyond the scope of this study. Although we are aware of the possible bias originating from it, we considered its effect on the final results marginal due to the large amount of data.

4. Methodology

As shown in Figure 1, our methodology comprises a sequential number of steps for pre-processing (defining temporal bins for before/during/after the Olympics and identifying residents vs visitors), textual analysis (sentiment analysis and automated semantic topic modeling), spatial hot spot detection, and finally evaluation and validation through a point pattern test of our results. The single steps are described in the following sub-sections.

4.1. Pre-Processing

We developed a two-step filtering procedure to prepare the raw data for the subsequent analysis:

Temporal binning: First, we created temporal bins from the raw data representing time periods before, during and after the Olympic Games (OG). This allows us to distinguish between "event days" and "comparison days". The reason for following this approach has been described in previous literature, as large-scale events such as the OG change the dynamics of a city for the time of the event. The temporal bins have been defined as follows: before: *June 27–July 13, 2012*; during: *July 27–August 12, 2012*; after: *August 27–September 12, 2012*.

Spatiotemporal subsetting (hypothesizing residents and visitors): The self-reported geolocation data from tweets and the frequency of their presence in the temporal subsets were used to identify presumable "residents" and "visitors" in London. Our approach is based on the work of Abbasi et al. (2015), who identified these user types in Sydney for city trip analysis. The rationale for identifying the two groups was the following: A person who tweeted at least once in each of the temporal subsets was considered a "resident", whereas a person who tweeted in just one of them was considered a "visitor" (non-resident). The remaining users of the dataset were not considered in the present study, as we could not differentiate between less actively tweeting residents or those visitors who stayed longer than a month, without performing further extensive analysis. Although, it is possible to identify them based on their tweets' content, but that is a complex methodology on its own and, therefore, was beyond the scope of this study. We are aware of the limitations of our method and discuss them, along with the advantages in the Discussion section. Yet, our results underpin that the method is effective when there is no additional data available for classifying user types, and it sufficiently reflects the necessary differences between the two groups for the desired purposes in our case.

4.2. Semantic Analysis

The semantic text analysis was performed in two consecutive steps: sentiment analysis, followed by automated topic extraction using the unsupervised machine-learning method Latent Dirichlet Allocation (LDA).

4.2.1. Sentiment Analysis

The sentiment analysis algorithm used in our approach is based on the work of Breen (2012). Sentiment scores were calculated for each tweet to automatically define to what degree it contains positive or negative sentiments, by calculating the difference between the number of positive words and the number of negative words.

This approach requires a dictionary with positive and negative words, for which we selected the Hu Liu lexicon (Hu & Liu, 2004), which is the most acknowledged dictionary in recent literature. Generally, if the score value is higher than zero, the sentence is assumed to contain an overall "positive sentiment", whereas it is considered containing a "negative sentiment" if the value is below zero. If the score equals zero, then the sentence is considered "neutral".

The main disadvantage of this approach is that the algorithm has limitations in defining unambiguous negative or positive scores for sentiment values around zero because they are indeed either neutral or they are misclassified with a comparatively high probability. Thus, we categorize positive tweets with the score equal or higher than 2 and negative tweets with the score equal or lower than −2. This does not mean that all the tweets with the sentiment value of 1 and −1 are neutral; rather they have a lower accuracy of being identified as positive

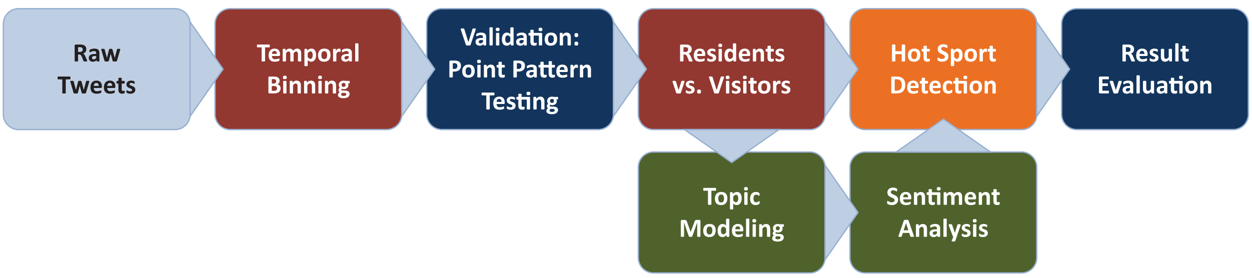


Figure 1. Workflow overview.

or negative and, therefore, we do not consider them in our analysis.

The terms "positive" and "negative" will be used throughout the article as defined above.

4.2.2. Machine-Learning Topic Modeling

As keyword-based approaches have limitation for social media data (Eisenstein, 2013), we used a machine-learning algorithm that extracts the latent structure of a dataset. This topic modeling approach clusters the data stream and filters the relevant tweets for further subsequent spatial analysis. Concretely, we used LDA, which is a probabilistic topic modeling algorithm that clusters semantic topics in a dataset. LDA is an unsupervised generative model that produces a document-topic distribution and a topic-word distribution. More information about the model and the hyperparameters of LDA can be found in (Blei, Ng, & Jordan, 2003).

Before the actual topic modeling procedure, social media posts need to be pre-processed, thus significantly improving the performance of LDA. We followed the steps defined by Resch et al. (2017), where every pre-processing step is explained in more detail. In the first step, every tweet is split at blank spaces so that every single character or sequence of characters can be treated individually (*tokenization*). Then all the words are set to lowercase to account for spelling mistakes and differences. In our experiment, URLs, special characters [e.g., ":" or ")"], short words (less than three characters), stop words (identified by a manual list and the list from Natural Language Toolkit (Manning et al., 2014, for English), and unique words that appear only once in the corpus, as well as numbers, are considered noise and are deleted. The remaining words are then reduced to their word stem using the Porter Stemmer (Porter, 1980).

In the next step, we applied LDA on the pre-processed data. We used the implementation of the Gensim library (Gensim, 2017) in Python and processed all the experiments with the following parameter values, which have been empirically derived, as no generically proven formal a-priori parameter estimation method exists so far: $\alpha = 0.0001$, $\beta = 1/\text{number\_of\_topics}$ and number_of_topics = 30. We set α to a value that is close to zero because short documents such as tweets usually only contain a single topic (Zhao et al., 2011). The other two variables were chosen according to experimental evidence. In the final step, we classified the tweets in accordance with the topic with the highest probability. The extracted topics were then manually interpreted, focusing on *Olympics*-related and *transportation*-related topics. From our perspective, a topic is related to transportation when words like London, station, railway, underground, etc. have a high probability in a topic, whereas a topic is considered *Olympics*-related if the stem "olymp" has the highest probability, and other words like stadium, ticket, wembley, athlete, etc. also have a high probability. Examples of Olympics- and transportation-related topics can be found in the Results section.

4.3. Spatiotemporal Data Processing

In order to study the spatiotemporal behavior of residents and visitors in the three temporal bins (before, during and after the OG), we analyzed daily and hourly tweet intensities for the subsets of positive and negative tweets and the main semantic topics (LDA output), as well as the similarity patterns for spatial point distribution (Figure 2). In the last step, we investigated spatial hot spots using Kernel Density Estimation (KDE). The maps illustrating the results of the KDE can be found in the supplementary file.

To quantify spatial similarity between tweets before, during and after the OG, a nonparametric and area-based spatial point pattern test was used (Andresen, 2009; Andresen & Malleson, 2013). The test requires the following datasets: base points and test points for comparing spatial patterns and base polygons representing the areal units. We had 5,888 polygons as areal units using the administrative dataset of the Greater London Lower Super Output Area (LSOA) from 2011 (Greater London Authority's DataStore, 2017). The LSOA areas are only used for the similarity test in our study, to define a general pattern in tweeting behavior. The base points are the tweets from the "during OG" bin, both for residents and for visitors in two consecutive analyses. Whereas the *test points* are the tweets posted before and after the OG, first for residents then for visitors. The entire analy-

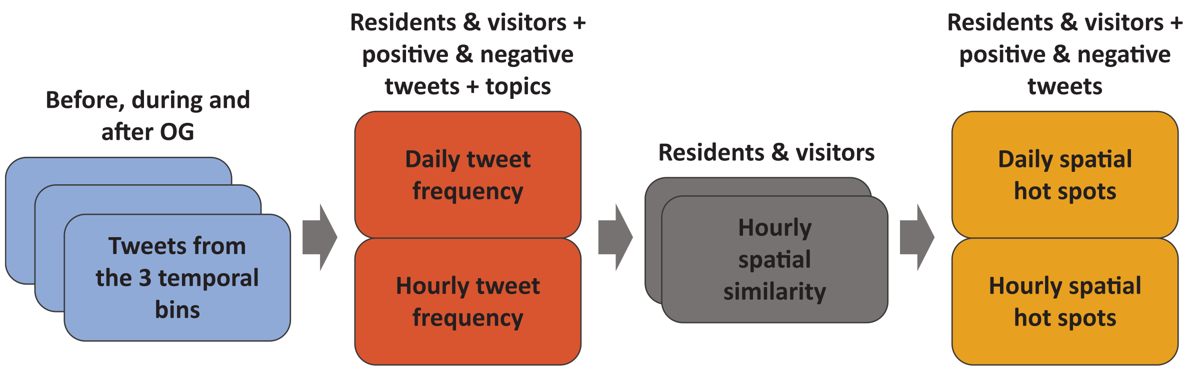


Figure 2. Spatiotemporal data processing overview.

sis was performed on hourly subsets. Regarding the base point dataset, the next step is to assign the points to the areal units (the LSOA) and then to calculate the percentage of points within each LSOA. For the dataset containing the test points, after assigning them to LSOA polygons as well, they should be randomly sampled, selecting 85% of the points and then calculate percentages (use Monte Carlo simulation to repeat this action 200 times). After that, we can create confidence intervals for each areal unit. Following these separate steps, the base percentage and test confidence for test interval are compared, and the result is the global index of similarity (for all the data) and the local one (for each areal unit). For this case study, the outcome of the test is a global index of similarity, where values range from 0 (no similarity) to 1 (identical). If the index is higher than 0.80, the two datasets are considered to be highly similar (Andresen, 2016). The index shows the level of similarity in the respective LSOA areas between the two analysis periods (Equation 1):

$$\frac{\sum_{i=1}^{n} s_i}{n} \quad \text{(Andresen, 2009)} \qquad (1)$$

where s_i is equal to one if two tweet datasets (in our case, similarity between the three temporal subsets, considered two by two) are similar in spatial unit i, and zero if the two are not similar at all. Further, n is the total number of spatial units (the LSOA polygons).

To visually analyze the spatiotemporal characteristics of our findings, we generated hourly and daily density maps for positive and negative tweets during the three time bins according to the user groups (residents vs visitors). There are many spatial tools used to understand changes in geographical patterns (Chainey & Ratcliffe, 2005). For this case study, we chose the KDE method, which involves placing a kernel over each observation (tweet), and, by summing these kernels, showing a density estimation of the observations' distribution (Fotheringham, Brunsdon, & Charlton, 2000). We chose KDE because it belongs to a non-parametric class of density estimators, which has no fixed structure and depends on the point data to define an estimate; practically the form of the density is determined only from the data without any model. The parametric methods, such as Maximum Likelihood Estimation or Bayesian Estimation, assume to know the shape of the distribution. In addition, KDE is highly used for frequency distributions allowing a quick exploration of the dataset distribution. In the article the bandwidth selection was performed automatically by the software used, ArcGIS 10.4, where the kernel function is based on the quadratic kernel function. One of the main advantages of KDE is that it determines the spread of positivism and negativism in this case study, namely the area around a cluster where the likelihood for a positive or negative polarity is present based on spatial dependency. First, we split the data into hourly and daily segments and then ran the nonparametric KDE tool for each layer and temporal bin, which helped to illustrate spatial changes in residents' and visitors' tweeting behavior.

5. Results

Table 1 shows the summary of our main results that were generated through the methodology described above to provide an overview of the content and structure in this section.

5.1. RQ1 & RQ2: Geolocated Tweet Density and Sentiment Intensity for Temporal Subsets

We defined three temporal subsets for our analysis: the time period of the OG and the same number of comparison days before and after the Olympics to test the effect of the OG on spatiotemporal tweeting behavior and on the tweets' semantic content relating to RQ1 and RQ2.

One essential step for this study was to identify presumable residents and visitors. By applying the criteria

Table 1. Results summary.

	Residents	*Visitors*
Positive	• August 4: high positive peak for residents (gold medals for Great Britain), hot spots in the city center and at the Olympic Park; however, no increase in residents' raw tweets intensity. • Opening Ceremony and Closing Ceremony clearly stand out in the number of positive tweets.	• August 4: positive sentiment peak in the daily temporal frame and slight increase in raw tweets intensity. • Well-defined spatial hot spot at the Olympic Park.
Negative	• Mostly flat distribution on daily temporal patterns. • More negative hot spots outside the city center during and after the OG.	• Low oscillations for all tweets and higher for the topics, e.g., the transportation topic. • Before the OG, negative sentiment exceeds the positive for a few days.
All tweets	• Residents and visitors show different temporal and spatial patterns. • Higher number of unique visitors tweeting during the OG. • Tweets' spatial distribution per hour shows the highest similarity during the night (low number of tweets) and low similarity during the morning and evening (high number of tweets).	

mentioned before (see Methodology section), we had to remove approximately 25% of the before OG tweets, 29% of the OG tweets and 24% of the after OG tweets. Practically speaking, we removed those users who tweeted exclusively in two temporal bins because it would have been difficult to distinguish whether they are just less actively tweeting residents or visitors who stayed a longer period than one month. The 11,571 London residents have the highest tweeting intensity during the OG compared to the visitors, who are texting more in the after OG period (Table 2). Regarding RQ1, we were able not only to distinguish residents and visitors in the dataset based on their temporal profile but also to identify clear and fundamental differences in the two groups' spatiotemporal behavior. Considering the different effect of planned events on residents and visitors, this finding has a key role in various planning-related social media analyses.

As for RQ2, large events tend to increase the social media participatory behavior (Wang, Can, Kazemzadeh, Bar, & Narayanan, 2012), which was also confirmed in this case by the highest density of tweets occurring during the OG (594,891 tweets). Another peak in tweeting intensity (545,693 tweets) was identified after the OG period (especially among visitors), which might be explained by the Paralympic games period and the London 2012 Festival as an accompanying event of the OG to organize "the most culturally engaging" OG in history (Brown, 2012).

Further, one of our hypotheses was that positive sentiments in the text will occur more often during a large event compared to other usual days for the same locations. This assumption was confirmed by the obtained sentiment scores for the six datasets used in this study: 7.65% of the resident tweets and 6.02% of the visitor's tweets during OG are positive, while just 3.04% respectively 2.24% are negative (Figure 3). There was a noticeable decrease in negativity, while the positivity increased.

5.2. RQ3: Semantic Topic Extraction

In every sub-dataset (spatiotemporally divided, see Methodology section), we can identify one or more related topics for our target topics "Olympics" and "transportation". Table 3 shows the ten words with the highest probability in the topic. Due to the limited space, we visualize only some of the topics. The reason for the missing syllables of the words is the pre-processing step, stemming, which cuts the word to its root.

Table 3 shows that we can clearly identify topics related to "Olympic" and "transportation", distinguishing the periods *before*, *during* and *after* the OG, as well as between residents and visitors. In all of the "Olympic"-

Table 2. Residents and visitors for the three temporal subsets.

	STEP 1		*STEP 2*							
	Total geo-tagged tweets		**Residents**		**Visitors**		**Residents %**		**Visitors %**	
	Tweets	*Users**	*Tweets*	*Users**	*Tweets*	*Users**	*Tweets*	*Users**	*Tweets*	*Users**
Before OG	478,551	46,357	195,319		160,922	22,851	40.82	24.96	33.63	49.29
OG	594,891	59,248	210,024	11,571	190,770	30,443	35.30	19.53	32.07	51.38
After OG	545,693	54,956	178,100		226,986	30,147	32.64	21.06	41.60	54.86

Note: * unique users.

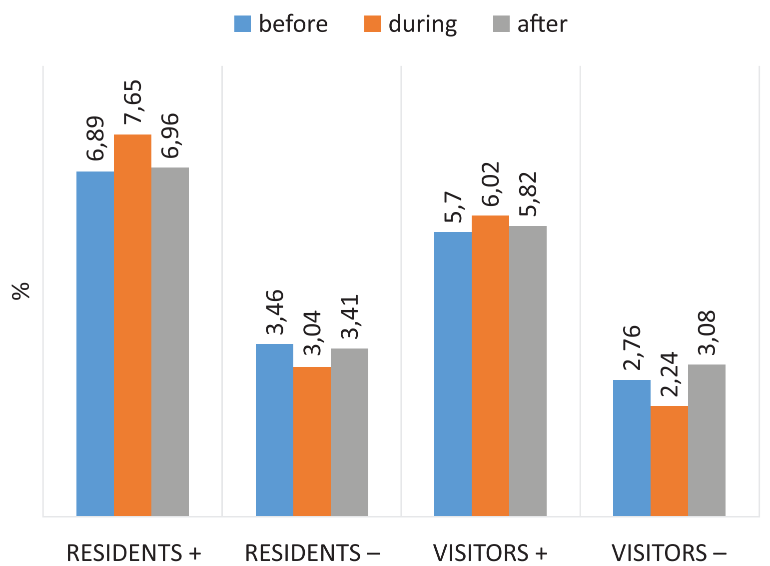


Figure 3. Sentiment analysis distribution for resident and visitors during the three temporal subsets (%).

Table 3. Examples of words and their probabilities for the identified topics "Olympic" and "transportation".

"transportation" visitors *before* OG		**"transportation" residents *during* OG**		**"transportation" visitors *after* OG**		**"Olympic" visitors *before* OG**		**"Olympic" residents *during* OG**		**"Olympic" visitors *after* OG**	
london	0.3995	london	0.3435	london	0.3768	olymp	0.0426	olymp	0.2676	paralymp	0.1133
greater	0.0938	greater	0.0798	greater	0.0692	loool	0.0310	excit	0.0400	olymp	0.0958
other	0.0707	other	0.0670	other	0.0587	year	0.0280	stadium	0.0395	stadium	0.0526
station	0.0412	station	0.0465	station	0.0370	walk	0.0238	final	0.0395	athlet	0.0194
hotel	0.0207	railway	0.0228	hotel	0.0205	point	0.0228	ticket	0.0385	found	0.0186
railway	0.0191	stratford	0.0202	bridg	0.0167	iphon	0.0219	photo	0.0362	serious	0.0158
underground	0.0106	street	0.0173	railway	0.0166	weird	0.0144	wembley	0.0240	watch	0.0156
victoria	0.0098	bridg	0.0160	tower	0.0122	shall	0.0133	game	0.0237	teamgb	0.0149
street	0.0094	venu	0.0142	arena	0.0116	togeth	0.0132	post	0.0233	paralympicsgb	0.0147
bridg	0.0074	underground	0.0116	cross	0.0103	restaur	0.0127	athlet	0.0188	problem	0.0139

related topics, the word "olymp" has a high probability and a significantly high probability during OG compared to the other words in the topic. In the case of the "transportation"-related topics, multiple words show high probabilities, such as "station", "railway" or "underground". It is notable that the same words in the "transportation"-related topic show similar probabilities in the datasets for residents and visitors in different time periods. In the dataset after OG, when the Paralympics took place, "paralymp" is also the most probable word in the "Olympic"-related topic.

5.3. RQ2: Similarity Index

Figure 4 shows the similarity values in hourly bins as defined above (see Methodology section). The highest similarity values occur during the night when tweets are posted from the same LSOA areas, but they don't have a high density, according to the hourly intensity results. Starting at 5:00 a.m. the similarity curve decreases until around 9:00 a.m. This shows that during the OG the spatiotemporal behavior of the users is different compared to before and after the OG. The more noticeable differences are at the end of the day, after 6:00 p.m., between residents and visitors after OG (ranging from 0.5814 to 0.5019), and between both visitor datasets (ranging from 0.5635 to 0.5019).

5.4. RQ1 & RQ2: Temporal Analysis

After extracting the topics and defining the sentiments of each tweet, we analyzed the temporal distribution of the negative and positive (see Methodology) tweets of residents and visitors, both on hourly and daily levels.

5.4.1. Daily Patterns

The daily tweet intensity using the raw Twitter data for *residents* showed two temporal peaks during the OG, at the Opening Ceremony and at the Closing Ceremony. The

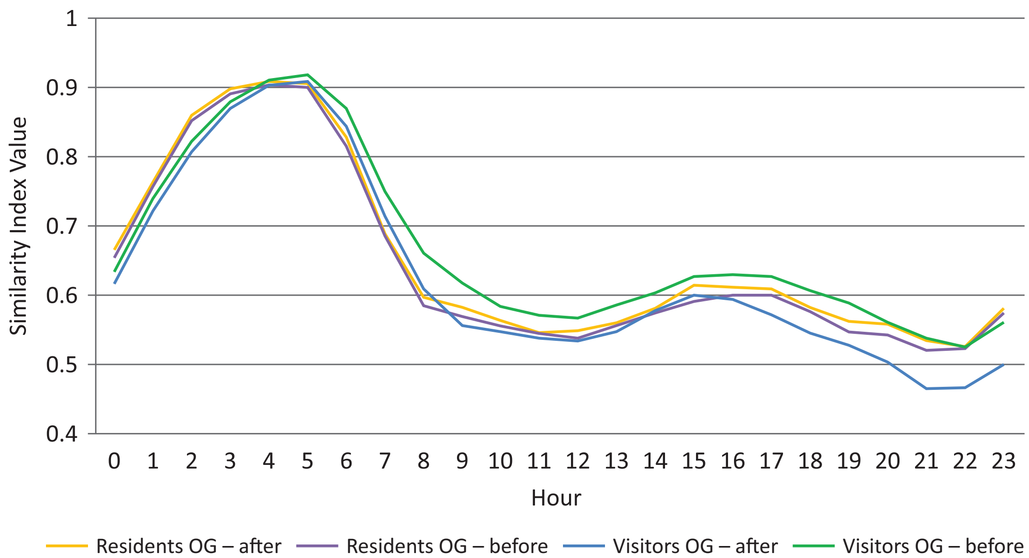


Figure 4. Similarity Index distribution for OG tweets as base points (range 0 to 1 where 1 means identical pattern for both analysis periods and 0 represents no similarity at all).

visitors' time series (unlike the *residents*) showed a peak in the OG period around August 4, when Great Britain won three gold medals in athletics. What is surprising is the higher volume of tweets after the OG for the visitors compared with the other time bins, including a peak during the Paralympics Closing Ceremony (Figure 5).

Next, we compared the daily patterns for sentiments in the data subsets. Figures 6–8 illustrate daily intensities in sentiment distribution for residents and visitors during the three temporal frames. It shall be mentioned that in a small number of days the tweets volume for the specific topics is low, especially for the "transportation" topic.

5.4.1.1. Positive vs Negative Tweet Trends

While analyzing the tweets for residents and visitors, the daily distribution of negative tweets was fairly equal and smooth for "all tweets" (all six data frames, all topics). The negative tweets for residents included in the "olympic" topic have an almost flat trajectory, similar to the ones including "all tweets" (Figure 7 vs Figure 6), except July 11, when they showed an increase and the hot spot map showed higher intensity in the London Center areas, Lewisham and Morden, close to Wimbledon. The intensity of positive tweets was more predominant than the negative ones at any time (Figure 6 and Figure 8), with higher values during the OG and spatial concentration around the Olympic Park (Figure 9). For the "olympic" topic, the positivity curve reaches its maximum for residents on August 4 (from 0.15% negative tweets to 0.86% positive tweets), while for "all tweets" the peak is higher for visitors (Figure 6). In the newspapers, this day is referred to as "Saturday night fever", when Jessica Ennis, Greg Rutherford and Mo Farah all

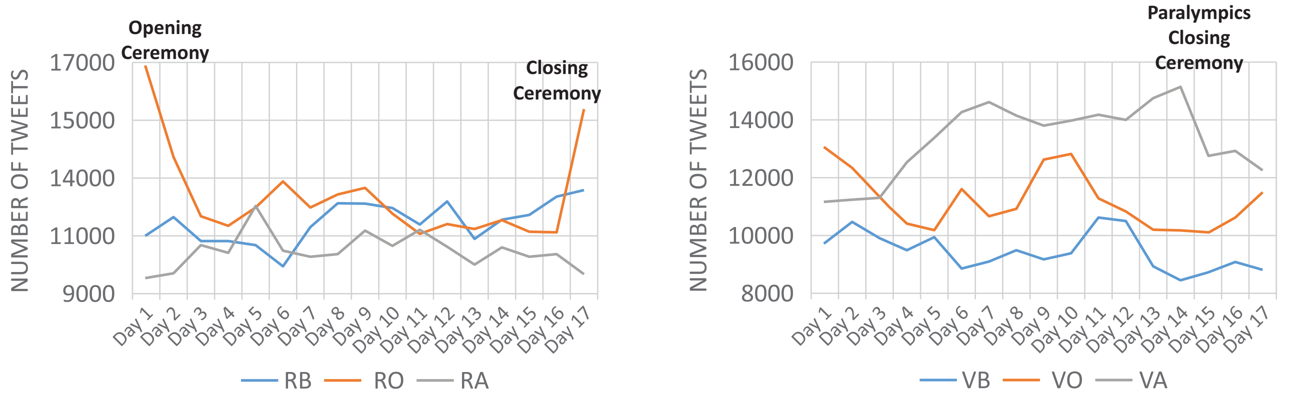


Figure 5. Daily tweets density for residents and visitors for the three temporal bins (R = residents, V = visitors; B = before OG, O = during OG, A = after OG).

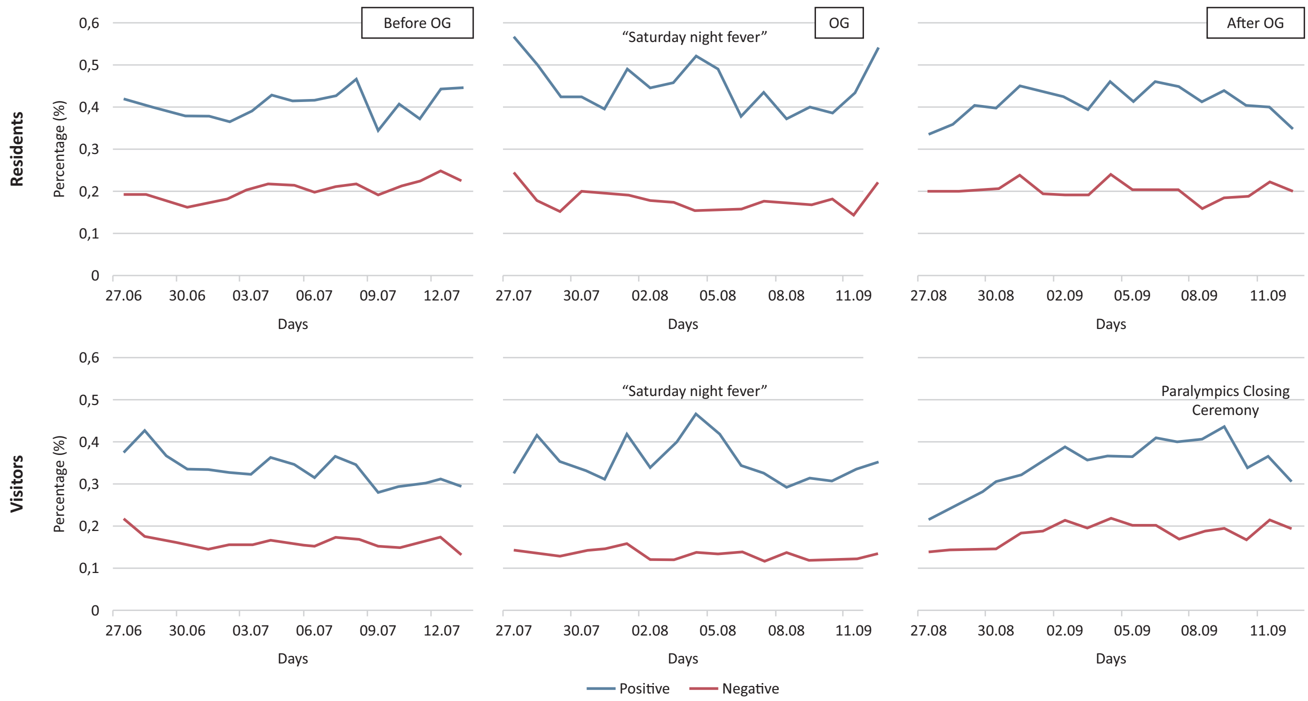


Figure 6. Daily sentiment analysis distribution for residents and visitors for "all tweets" (percentage of total number of tweets in the given temporal bin for the respective categories).

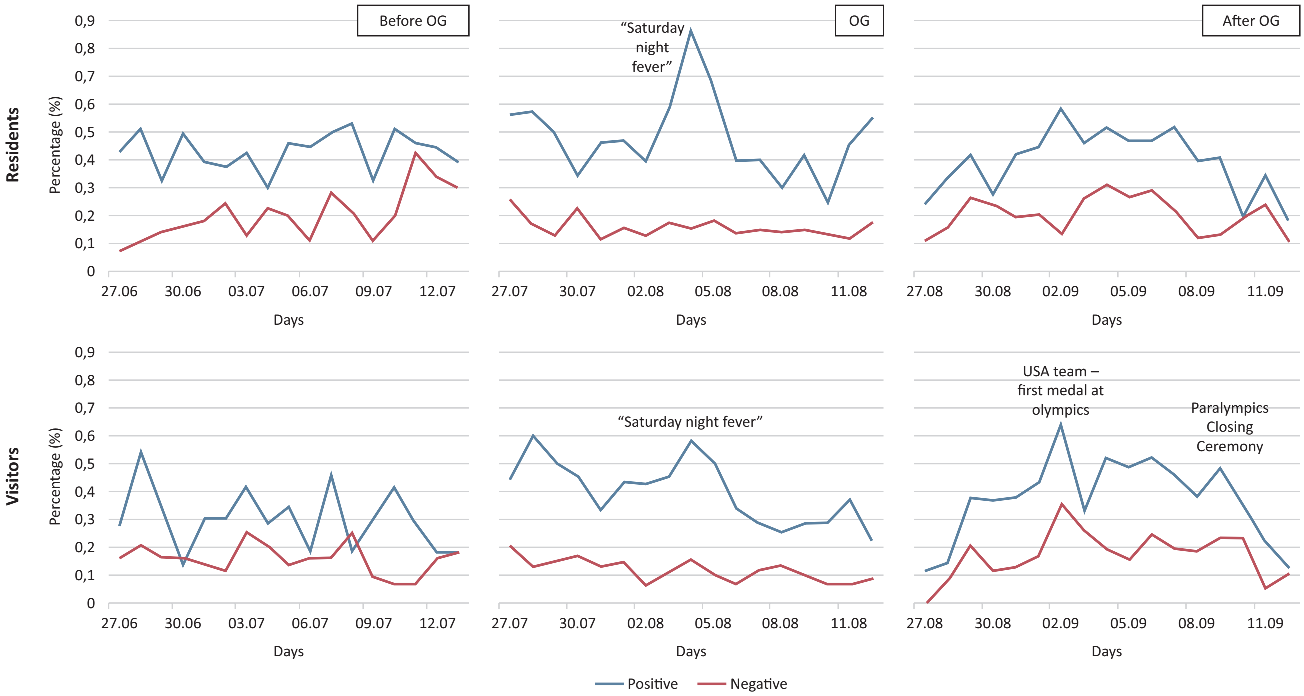


Figure 7. Daily sentiment analysis distribution for residents and visitors for the "olympic" topic (percentage of total number of tweets in the given temporal bin for the respective categories).

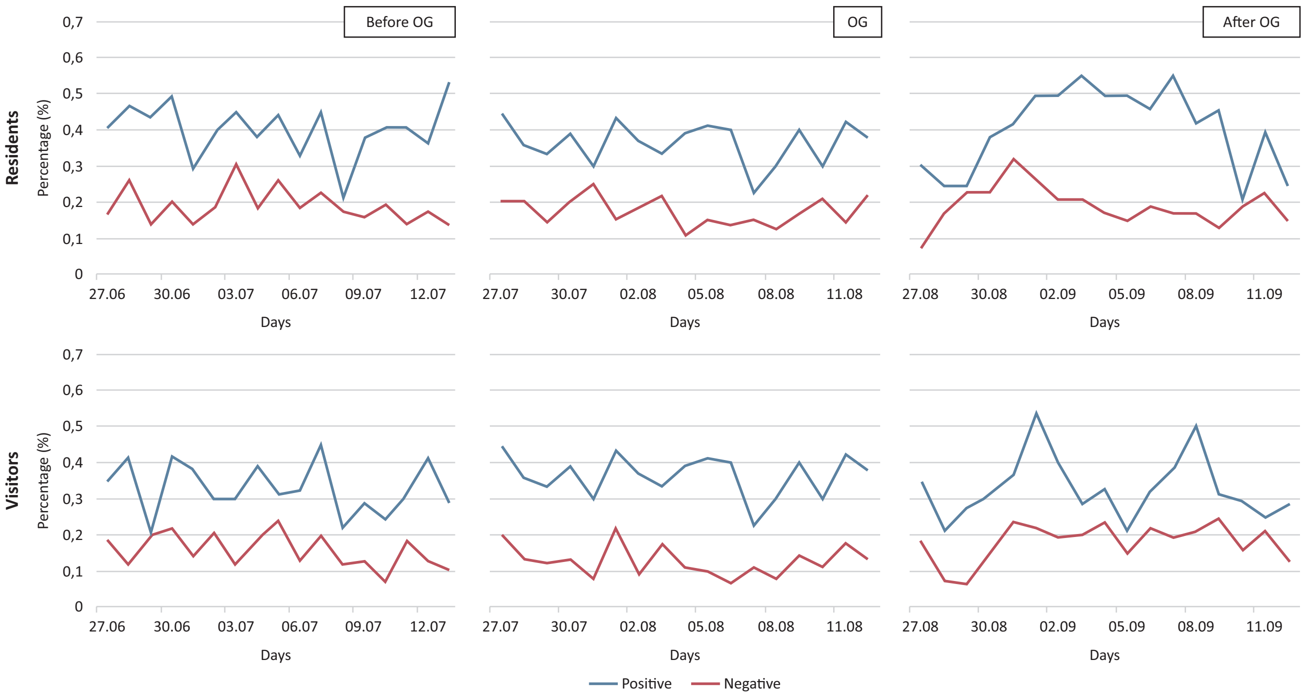


Figure 8. Daily sentiment analysis distribution for residents and visitors for the "transportation" topic (percentage of total number of tweets in the given temporal bin for the respective categories).

won gold medals for the host nation. This shows higher public engagement when an action such as winning a prize by co-nationals takes place.

Regarding the "after OG" period, on September 2, when the USA team won the first medal in the trunk and arms mixed at the Paralympics, a decrease of negativity happened for residents, while during the same day an increase occurred for the visitors, together with an increase in the positive tweets (Figure 7). A common positive peak can be observed for "all tweets" and "olympic" on September 9 (Figure 6 and Figure 7), during the Paralympics Closing Ceremony. In comparison, the sentiments distribution of the "transportation" topic contains a smoothed zig-zag line, and an increase followed by a decrease in the positive tweets after the OG. Interestingly, on August 4 there was no peak in the results for either residents or visitors, compared to the other subsets. The tweeting behavior after the OG for "transportation" shows a higher difference between positive and negative tweets for residents, mostly from September 1 to September 9 (Figure 8). However, the maximum tweet volume is 30 per day.

5.4.1.2. Daily Trends of Residents vs Visitors

Before the OG, residents and visitors for "all tweets" showed slight changes in the trend line (Figure 6), while for the topic "olympic" the visitors' tweets tended to form a zig-zag-like time series (Figure 7), similar to the "transportation" topic (Figure 8). No daily spatial hot spots were found in the Queen Elizabeth Olympic Park for this period.

During the OG, residents and visitors for "all tweets" showed a higher volume of positive tweets, on August 1 and August 4 (Figure 6). On August 1, the spatial hot spots are distributed between London's central area and the Olympic Park area, and on August 4 a high density is located specifically around the Olympic venues: the Olympic Park zone, the River zone (including Greenwich park), and the Central zone (including Hyde Park and Regent's Park). For the residents, we also notice smaller hot spots in many parts of the city, which suggest an increased interest in people's tweeting behavior for a special occasion (Figure 9). August 4 is also a common peak for "olympic" tweets, mostly for residents, with a hot spot location around the Olympic Park, the city center and another one between these two as well, almost continuously. In the same time, the visitors show the hot spot only around the Olympic Park and with much lower intensity in the city center (Figure 7, 9). Interestingly, August 4 showed a positive peak that is not connected with the increase in the intensity of the raw tweets for residents. Another dissimilarity arose for the "transportation" topic, where the graphic of sentiments distribution showed a different pattern (Figure 8). For example, the highest positive peaks for the OG period occur on August 1 for the visitors and August 2 for the residents. The visitors' tweeting hot spots are in the city center and at the Olympic Park, while the residents' tweets are clustered in an elongated hot spot with median values around the Olympic Park.

For the *After OG* period, September 2 was a peak of positive emotions for residents and visitors for the

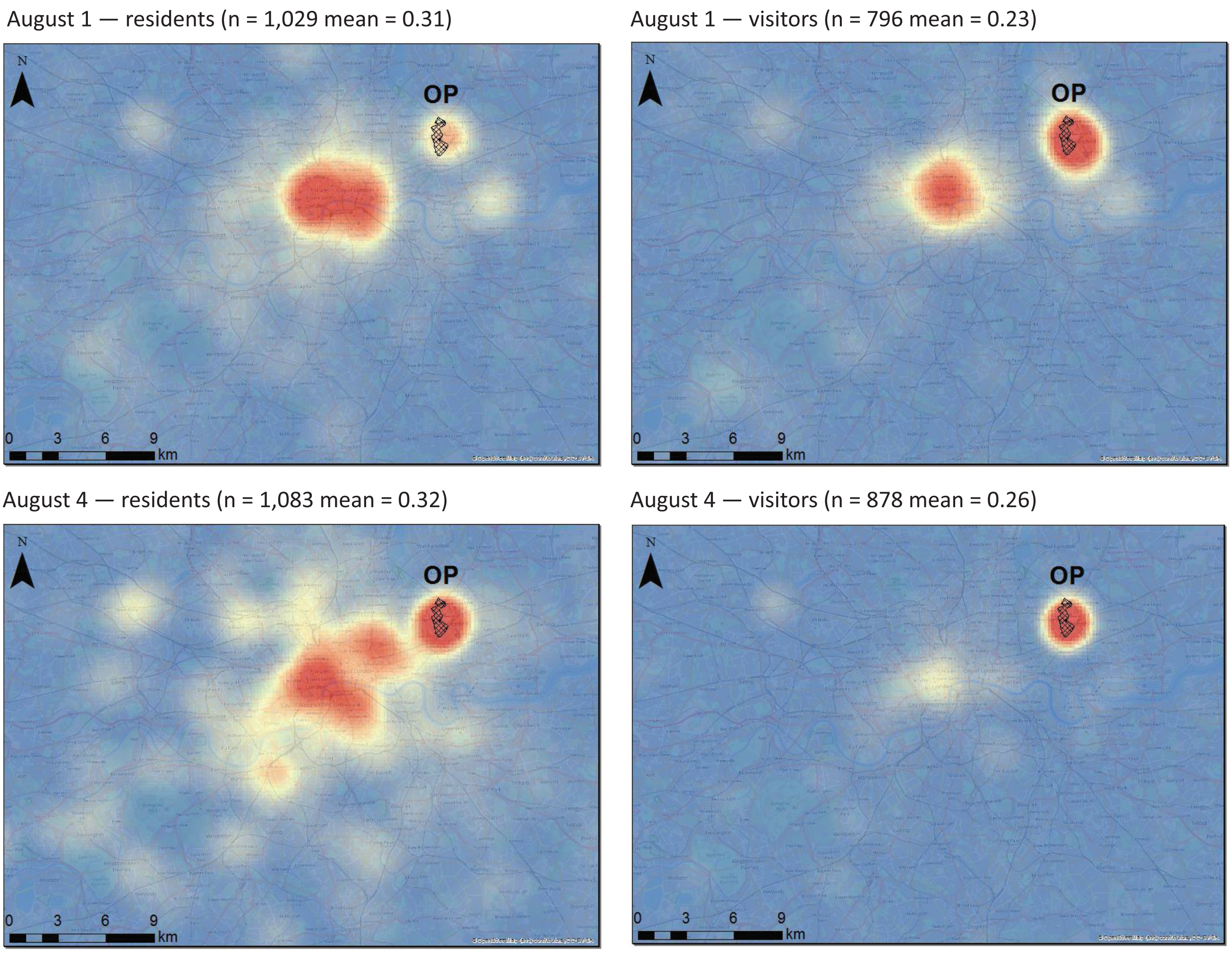


Figure 9. Hot spots of positive tweets (OP = Queen Elizabeth Olympic Park; red = high density, yellow = average density, blue = no density).

"olympic" topic, showing an intense hot spot for the visitors at the Olympic Park, while the residents' hot spot included the park, but the center was shifted towards the western part of the park. On September 9 residents from the "olympic" topic (Figure 7) and visitors for this topic and "all tweets" (Figure 6) showed an increased positive feeling, possibly caused by the Paralympics Closing Ceremony. Also, on the spatial density for this day the hot spots are located in the approximate city center and in the Olympic Park zone. A different temporal pattern occurred for the "transportation" topic of the visitors' tweets (Figure 8). Two predominant days showed positive peaks, including September 1, when the majority of tweets are from the Olympic Park, and September 8 when all the active tweeting happened in the London center.

5.4.2. Hourly Patterns

After identifying peaks and patterns in the data on a daily level, we also analyzed the tweets' hourly spatial and temporal distribution. Figure 10 shows a general overview of the hourly distribution for the raw number of tweets per user groups in the two weeks temporal frames (a total of six weeks). Residents before OG showed a cyclical daily circular pattern, with low intensity overnight then a rapid increase in the morning and a steep decrease a few hours after midnight. Interestingly, we can define peaks for residents and visitors during the OG for the Opening Ceremony and also for the Closing Ceremony, together with the "Saturday night fever" mentioned in the daily patterns on August 4 for residents. Another peak in tweeting intensity occurs on September 9 for the residents, possibly due to the Paralympics Closing Ceremony.

After summing up the number of tweets per hour, we noticed only a slight change in the temporal tweeting behavior of the residents between the OG period and comparison days. However, after hour 20, a small increase occurs when compared to the "before" and "after" data (Figure 11). For the visitors, the tweets volume is surprisingly higher during the afternoon and the evening considering the after OG period. The number of unique users after the OG was lower than during the OG, which shows that there were fewer users after the OG, but they were more active than the ones who tweeted during the

Figure 10. Residents and visitors tweeting behavior per hour for the three temporal bins.

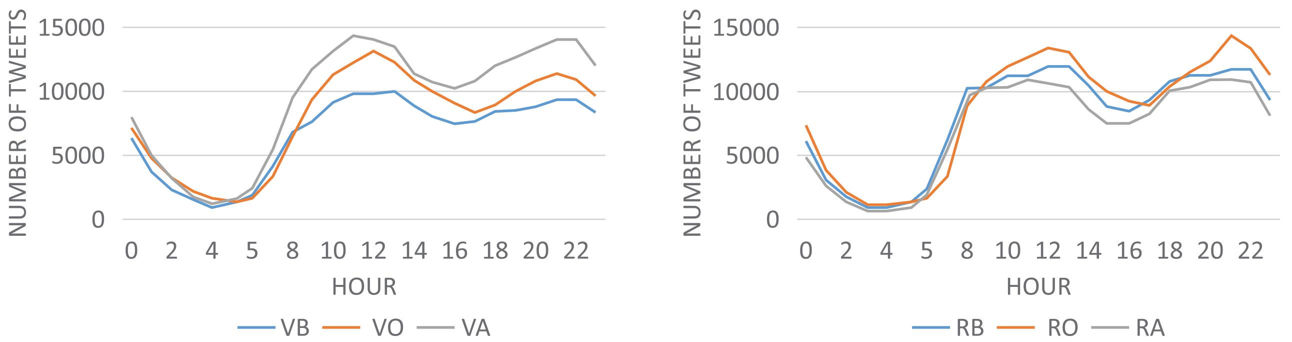


Figure 11. Hourly distribution of the tweets during the three analysis periods (R = residents, V = visitors; B = before OG, O = during OG, A = after OG).

OG. This observation may be explained by various factors, such as the start of school when students are more active, who may potentially be misclassified as visitors because of their limited Twitter activity over the summer from the school/university location (in London).

Furthermore, the hourly patterns were analyzed considering the sentiment score and also the topic allocation. Figure 12 shows the amount of positive and negative tweets for both groups per hour. It is evident that the number of negative tweets never exceeds the number of positive ones in either of the user groups. The positive peaks differ between residents and visitors: Residents tended to tweet more in the evening (around 9 p.m.), whereas visitors tweeted slightly more around 12:00 noon. If we only consider the tweets where the topic "olympic" was identified by the algorithm (Figure 13), these characteristics become even more striking. However, if we compare the relative tendencies (all positive/negative tweets in the given hour for each period) these peaks are smoother, which means, there are generally higher number of positive tweets in the evening.

The following videos[1] show the spatiotemporal pattern of positive and negative tweets (aggregated to 24h hours such as for Figure 12, in 10-minutes timeframes) in both user groups for all the three analysis periods. (A static version containing four different hours during the day can be found in the supplementary file.) Blue points represent negative tweets, whereas the positive ones are visualized in red. The semi-transparent points representing each tweet stay there for two hours to illustrate the density of tweets. Each video shows all the three temporal bins after each other (3 × 24 hours) for our two groups of users (residents and visitors), and a clock shows the current time in the lower right corner.

5.4.2.1. Changes in the Pattern Comparing the OG Period to Before and After

This section reflects to RQ2, as we compared the before and after OG periods to the patterns during the OG. For the residents, we can see that the core of the main hot spot is constant for each hour throughout all the analysis periods, and it is located mainly in the city center. For the negative tweets, the before and after periods are quite similar during the day, but for the positive tweets, there is still a hot spot around the Olympic Park. The reason for this is that our analysis period after the OG includes the days of the Paralympic Games as well. At the same

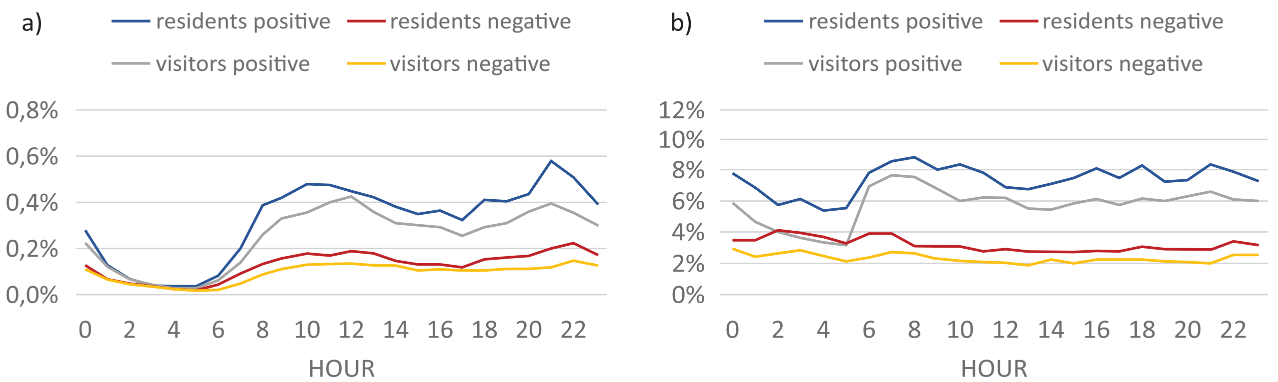
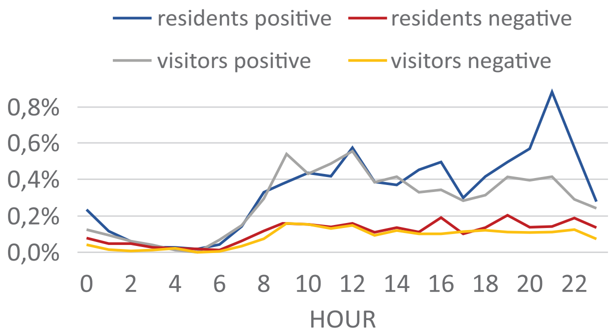


Figure 12. Positive and negative tweets per hour (%): a) as *absolute* values–every hour compared to the number of all tweets during the OG for residents and for visitors; b) as *relative* values–every hour compared to the number of all tweets in that hour during the OG for residents and for visitors).

[1] http://giscience.zgis.at/gisce/Videos_Towards_Citizen-Contributed_Urban_Planning_through_Opinion_Mining_of_Twitter_Data

time, for the visitors, the pattern during the OG and in the other two periods is not that different. Except during the morning hours, but the lower number of tweets can explain this because in this case even 2–3 point can result in relatively strong hot spots.

Figure 13. Positive and negative tweets per hour (%) for residents and visitors for the "olympic" topic.



5.4.2.2. Pattern of the Average Day: Before, During and After the OG

For residents during the OG, there are a few extra hot spots reflecting the venues of the OG. In the periods before and after the OG, the smaller hot spots in the outer parts of the city occur mostly in the morning and the evening, probably due to commuting, but only for the negative tweets. Interestingly, for positive tweets from visitors, the morning hot spot is more concentrated (they do not commute) but only before and during the Olympic Games. During the rest of the day, the patterns do not change significantly in any of the analysis periods, both for positive and negative tweets.

5.4.2.3. Positive vs Negative Tweet Trends

In general, the positive tweets of the residents tend to be more concentrated with one main hot spot, except the morning after the OG where the negative tweets are conglomerated. However, during the OG in the evening, the size of the positive tweet hot spot is much larger. Probably that was the time when most of the residents were tweeting about the Olympics. For visitors, this concentration of one large hot spot for the positive tweets is not significant; there are smaller hot spots around the city in both cases, and, in general, the hot spot for the negative tweets is larger in extent.

5.4.2.4. Residents vs Visitors

The most significant difference in patterns between residents and visitors is the distribution of negative tweets in the morning, but again this could be a result of the low number of tweets. Also, the smaller hot spots of the visitors' tweets tend to be more on the Eastern side of the study area, especially in the period after the OG. In general, it can be significant for planners to further investigate the trends and possible causes for negative tweets, as partially it might be connected to planning-related issues such as low satisfaction with infrastructure or poor quality of services.

6. Discussion

6.1. Integration of the Results into Planning Processes

The major objective of our case study was to illustrate the general potential of Twitter data analysis for urban planning purposes in the case of large planned events. Thus, we identified and addressed research gaps, such as the distinction between residents and visitors regarding the Olympics and comparing event days and non-event days along with both spatiotemporal and content analysis in one study. Consequently, our results serve as a basis for further, more in-depth analyses.

In general, both previous research and the work presented in this article have shown that results from social media analysis are directly usable in urban planning processes, including the general ability to detect sentiments that are associated with places (Resch et al., 2016). In this regard, social media provide people (local citizens and visitors) with a simple and powerful instrument to share their opinions and subjective impressions. This is particularly relevant with respect to connecting social media posts to specific urban events such as Olympic Games or other large sports events, for gaining insight into the perceptions of the urban population regarding these events. In fact, social media are a valuable, open source of information for urban planning.

This openness is of particular importance because urban planning processes are oftentimes still characterized by closed communication between local and official actors, lacking open discussion and transparent procedures (Resch et al., 2016). Moreover, openness and transparency are increasingly a key factor for successful urban planning, allowing for an efficient weighing process that considers the opinions and sentiments of different stakeholders. Current planning processes, however, are mostly shaped by deductive processes, which are typically introduced and controlled by urban governments, oftentimes neglecting or not sufficiently integrating the needs of the citizens. In this context, social media play a key role because they provide an instrument for organizing public participation activities and citizen initiatives. On the positive side, the integration of public discussions on social media and other digital platforms also increase the validity and acceptance of governmental decision-making because traditional planning methods are complemented by new "human sensor" data that reflect the citizens' wishes and needs (Zeile, Resch, Exner, & Sagl, 2015).

This is in clear contrast to top-down approaches that follow different decision-making principles. Integrating social media into urban planning may be able to provide unseen insights into citizens' thoughts, perceptions and expectations concerning urban events in an inductive bottom-up approach. In this sense, urban planning-related discussions are, to some degree, self-organizing, giving citizens the chance to discuss planning issues in a peer-to-peer process, rather than in a government-driven one. However, the issues of the digital divide, that mostly younger, better educated, and more technologically savvy people participate in social media networks, should be addressed in social media analysis (Czepkiewicz, Jankowski, & Młodkowski, 2017). Due to this digital divide social media platforms are currently not representing the entire society or population appropriately (Diaz, Gamon, Hofman, Kiciman, & Rothschild, 2016; Mellon & Prosser, 2017), therefore conclusions drawn from the analysis depending on the phenomenon should be handled accordingly. The extremes are especially underrepresented in terms of age (very young, and older generations), economic situation (those who cannot afford access through internet or gadget), etc. Consequently, we are aware that the social media-based approach shows a number of limitations; still it may complement current urban planning procedures through an improved understanding of the city as a living organism through proactively engaging citizens into urban planning (Resch, 2013).

Based on the methods we used and their outcomes, we can identify two main types of further planning-related investigations (the list of the examples is not complete):

a) *Macro-scale:*

- On a city level, it is possible to point out the differences in the general mobility patterns compared to non-event days (also at different times during the day) and use it for further transportation modeling. The different needs of residents and visitors should be considered;
- Planners can also further investigate the hot spots for negative tweets in both groups (residents and visitors); as they show different trends, there might be different reasons behind them. These hot spots can be compared with the extracted topics in these areas, whether they are related to transportation or other planning-related topics, the event itself, or something else entirely;
- Regarding the extracted topics, it is possible to search for more specific terms, if the planners provide expert knowledge. Furthermore, other terms can be identified that are connected to the planning-related topics and have not been considered by urban planners yet. The sentiment and spatial distribution of these topics all over the city can then be explored, both for residents and visitors, as these might also differ in this case.

b) *Micro-scale*:

- We could clearly identify activity patterns related to individual venues of the Olympic Games. An interesting example could be to focus on a venue and explore the behavioral patterns of residents and visitors and the effect of a given event at that venue. (Specifically, right before or after and during the analysis.) Do the residents tend to be more negative? Or maybe less active during that time?
- Additional datasets are definitely advantageous for the micro-scale analysis. Planners can explore deeper connections between the event and other urban processes. For example, the effect on the local economy of using bank cards can also be analyzed. (Habidatum, 2017);
- Extracting information on a user level is also an option. For visitors, planners can trace the intra-urban mobility patterns, if they tweet regularly during the day (between their accommodation and the venues). This analysis is even more accurate with additional mobile phone data analysis.

6.2. Psychological Biases in Human Language and Social Media and Their Relevance for Urban Planning

Most generally, there is a universal positive bias in human language (Dodds et al., 2015). The findings of the present study are congruent with this kind of bias: A higher percentage of positive tweets than negative tweets were identified. Moreover, the residents show a clear peak of positive sentiments during the OG, even though there have been several examples observed recently where local people opposed the organization of the Olympics in their cities (e.g., Kaufmann, 2015; Moore, 2015; Sims, 2017). It might suggest that, once underway, world-class events in a city boost self-respect and pride of the city residents, and the perceptions of their benefits are typically optimistic (Whitson & Macintosh, 1993), which might not be true in the planning phase (e.g., Dempsey, & Zimbalist, 2017). However, it is important to consider that the positive cognitive bias and the homeostatic happiness maintain satisfaction in life, and self-beliefs can act as reality buffers (Cummins & Nistico, 2002).

This, again, raises the question of whether the positive cognitive bias can be a buffer that masks some inconveniences in the city that large events such as the OG may cause. For example, Ritchie, Shipway and Cleeve (2009) identified that, in general, urban residents supported important events in their area, but were concerned with some issues such as traffic congestion and an increasing cost of services. Additionally, the benefits of large events in a city can differ between social groups.

For example, younger residents, residents that have a higher socioeconomic status, and residents that live farther away from the event's location, are more likely to perceive additional benefits from the event (Ritchie et al., 2009; Whitson & Macintosh, 1993). The results of our study suggest positivism related to the OG, but beyond this social media positivism, there are several considerations explained above that urban governments and urban planners need to study. In other words, the obtained results are good indicators of the importance of large sports events for residents' life satisfaction, but, in urban planning, these results cannot be isolated from the rest of the city dynamics.

These issues become critical if we consider that urban governments are usually open to investing in consumer and entertainment-oriented developments, such as sports events (Harvey, 1987). However, citizens in a city are more than consumers. Additionally, a large sports event causes changes in different dimensions in a city such as image, knowledge, and emotions, where the long-term effects of these changes are complex to understand (Preuss, 2007). We believe that long-term social media analysis can be considered a necessary instrument to monitor these effects in a city and to offer more pluralistic information to urban planners. Urban planners can use this information to evaluate different qualitative and quantitative costs and benefits of a large sports event. Further research needs to develop new approaches to study large urban events' legacies using social media. At the same time, these approaches need to be enriched with robust epistemologies to understand the complex and dynamic human behavior in the virtual world (social media), without disconnection of the human behavior in the real world.

6.3. The Effect of the Paralympic Games on the Selection of the Temporal Bins

We were aware of the fact that there were days related to the Paralympics in the third temporal bin and, thereby, our comparison might show less significant differences. However, we also tested a fourth temporal bin (September 27–October 13, 2012), and the patterns in the original temporal bin (August 27–September 12, 2012) were not biased, except the day of the closing event. Therefore, we decided to keep the original after OG temporal bin because selecting days so much later can also have an effect on the final results, and those would be more difficult to interpret, such as different seasonal effects, or extraordinary events.

6.4. Identifying Residents vs Visitors

The process of categorizing Twitter users into "residents" and "visitors" is challenging, and, to our best knowledge, there is no "ground truth" methodology, providing unquestionable results, in the related literature. Also, Abbasi et al. (2015) stated that dividing social media users into residents and visitors is not an easy task. They categorized these types of people in Sydney for city trips supporting urban planning. Our study follows an adapted method from the original when the residents are defined as users tweeting at least ten times in at least $n-1$ phases of the temporal data analysis. One reason for using this adapted approach is that our datasets had specific time frames related to the OG event, and we hypothesized that for finding so-called active residents, they would need to tweet in all n temporal bins, while the visitors had to tweet just in one of the temporal bins.

The spatiotemporal patterns identified and described in this study show a relative verification of this approach. For example, August 4 was an important day for Great Britain and it was undoubtedly reflected in the spatial and temporal patterns for both presumable residents and visitors: as a monocentric, well-defined daily hot spot around Olympic Park for the visitors, and as polycentric hot spots for residents in many parts of the city, with particularly high density around the city center and also around the park. The daily temporal graphic for positive tweets supports the hot spot map for residents by highlighting a larger increase in positive sentiments compared to visitors. This may be because visitors are generally more excited and tweeting positively for all the OG results, while residents are more interested in Great Britain's performance. August 4 is important in the analysis because it emphasizes people's behavior and how the positive event of winning three gold medals changes the spatial distribution of tweets. Also, for the hourly hot spot detection, we notice more intense hot spots around the Olympic Park for visitors than for residents, i.e. at 8:00 a.m. or 12:00 noon there are only mild or non-existent hot spots for residents at the park.

However, this particular approach has limitations: We did not consider the declared language of the users (e.g., maybe non-English speakers are more likely to be visitors tweeting before and during OG); the tweets' "user location" field was not used considering the biased information introduced subjectively by the user; data availability–we only had access to 2012 London data, whereas having worldwide data would have been helpful for exploring user's activity status. An interesting future approach may be to adopt all location-related features and create an index of defining residents and visitors.

6.5. Topic Modeling

In our analysis we used the basic LDA model for topic modeling that follows a "bag of words" approach, meaning that it uses solely the frequency of terms in a document and does not take grammar or word order into account. A significant problem is names, which consist of two words like "Greater London", where each word is treated independently. As "Greater" and "London" are words that are commonly used in combination, analyzing biterms may increase the quality of our results. However, there are other names like "Olympic Games" where

"Games" is a common word in an English conversation. For our particular case, relevant biterms include "Greater London", "Olympic Games" and "Victoria Station", which ware included as single words in Table 3.

6.6. Influence of Topic Modeling on the Sentiment Score

The distribution of positive tweets during the OG for the "olympic" topic is different for both residents and visitors. While depicting the possible reasons, we noticed the different word probabilities resulting from the LDA algorithm, such as the word "olymp", with a probability of 0.2676 during the OG period for residents and 0.0426 after the OG for visitors. The inclusions of words not relevant to the topic (which is subjectively named after checking the highest words probabilities), such as "iphon" in the Olympic topic might lead to an unexpected temporal distribution. Also, the sentiment score function is limited in defining a high volume of positive and negative tweets, their majority being labeled as neutral, which may result in increased fuzziness in interpreting the results.

7. Conclusion

In conclusion, our findings validly answer our research questions: Through spatiotemporal and sentiment analysis of the tweets, we could identify significant patterns in terms of our two defined user groups as well as for the days before, during and after the event. Additionally, the uncertainty originating from the identification of the members of each group due to the lack of additional details can be reduced by integrating further datasets (e.g., cab rides, bicycle network and public transport usage, mobile phone data).

Regarding the utilization for planning purposes, we can state that despite the limitations described above, by applying our workflow to the sample dataset we can provide valuable information about the spatiotemporal behavior and sentiment of residents or visitors concerning large planned events. By comparing our results to important dates of the event (e.g., the Closing Ceremony, "Saturday night fever") or location of the venues, we could validate our results both content-wise and for spatiotemporal patterns, even on finer spatial and temporal scales. Last but not least, topics that are directly related to planning and transportation could be extracted and can be further analyzed for specific urban planning purposes in the future.

Concluding, the case study was also appropriate to illustrate the potential of utilizing social media data for sentiment analysis and topic modeling in order to provide general feedback regarding large planned events. Nevertheless, there are possible ways for improvement beyond the scope of the current study that can also aid to overcome some of the already mentioned limitations. One such option is to design a geovisual analytical tool to interpret the large amounts of data (e.g., maps, graphs, tweets, time periods), also supporting users who are less familiar with GIS concepts and methods. Furthermore, as an outlook to participatory planning, the acquired knowledge could be presented in a Volunteered Geographic Information platform, which is directly connected to the event and where people can provide feedback with location data.

Acknowledgments

This research was funded by the Austrian Science Fund (FWF) through the Doctoral College GIScience at the University of Salzburg (DK W 1237-N23). We would also like to express our gratitude to the FWF for supporting the project "Urban Emotions", reference number I-3022.

Conflict of Interests

The authors declare no conflict of interests.

References

Abbasi, A., Rashidi, T. H., Maghrebi, M., & Waller, S. T. (2015). Utilising location based social media in travel survey methods: Bringing Twitter data into the play. In *Proceedings of the 8th ACM SIGSPATIAL International Workshop on Location-Based Social Networks* (pp. 1–9). Redondo Beach, USA: ACM.

Andresen, M. A. (2009). Testing for similarity in area-based spatial patterns: A nonparametric Monte Carlo approach. *Applied Geography*, *29*(3), 333–345.

Andresen, M. A., & Malleson, N. (2013). Spatial heterogeneity in crime analysis. In M. Leitner (Ed.), *Crime modeling and mapping using geospatial technologies* (pp. 3–23). Dordrecht, Netherlands: Springer.

Andresen, M. A. (2016). An area-based nonparametric spatial point pattern test: The test, its applications, and the future. *Methodological Innovations*, *9*, 1–11.

Aubrecht, C., Ungar, J., & Freire, S. (2011). Exploring the potential of volunteered geographic information for modeling spatio-temporal characteristics of urban population A case study for Lisbon Metro using Foursquare check-in data. In *Proceedings of the 7th International Conference on Virtual Cities and Territory* (pp. 57–60). Lisbon: Nova University of Lisbon.

Blei, D., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993–1022.

Bollen, J., Mao, H., & Pepe, A. (2011). Modeling public mood and emotion: Twitter sentiment and socioeconomic phenomena. In *Proceedings of the Fifth International Conference on Weblogs and Social Media* (pp. 450–453) Barcelona, Spain: AAAI Press.

Breen, J. O. (2012). Mining Twitter for airline consumer sentiment. In G. Miner, J. Elder, A. Fast, T. Hill, R. Nisbet, & D. Delen (Eds.), *Practical text mining and statistical analysis for non-structured text data applications* (pp. 133–149). Waltham, MA: Elsevier Academic Press.

Brown, M. (2012). Cultural Olympiad 2012 reaches the critical masses. *The Guardian*. Retrieved from https://www.theguardian.com/culture/2012/mar/12/cultural-olympiad-2012-critical-masses

Castells, M. (1996). *The rise of the network society: Volume I: The information age: Economy, society, and culture* (1st ed.). Oxford, UK: Blackwell Publishing.

Cebelak, M. (2013). *Location-based social networking data: Doubly-constrained gravity model origin-destination estimation of the urban travel demand for Austin, TX* (Master Thesis). Faculty of Graduate School, University of Texas, Austin.

Chainey, S., & Ratcliffe, J. (2005). *GIS and crime mapping* (1st ed.). Chichester, UK: John Wiley & Sons.

Chen, Y. (2012). *Urban strategies and post-event legacy: The case of summer Olympic cities*. Paper presented at the AESOP (Association of European Schools of Planning) 26th Annual Congress, Ankara, Turkey, 11–15 July 2012.

Cheng, Z., Caverlee, J., & Lee, K. (2010). You are where you tweet: A content-based approach to geo-locating Twitter users. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management* (pp. 759–768). Toronto, Canada: ACM.

Collins, C., Hasan, S., & Ukkusuri, S. V. (2013). A novel transit rider satisfaction metric: Rider sentiments measured from online social media data. *Journal of Public Transportation*, *16*(2), 21–45.

Cook, I. R., & Ward, K. (2011). Trans-urban networks of learning, mega events and policy tourism: The case of Manchester's Commonwealth and Olympic Games projects. *Urban Studies*, *48*(12), 2519–2535.

Corney, D., Martin, C., & Göker, A. (2014). Spot the ball: Detecting sports events on Twitter. In *Proceedings of Advances in Information Retrieval* (pp. 449–454). Amsterdam, Netherlands: Springer.

Costanza, R., Fisher, B., Ali, S., Beer, C., Bond, L., Boumans, R., . . . Snapp, R. (2007). Quality of life: An approach integrating opportunities, human needs, and subjective well-being. *Ecological Economics*, *61*(2/3), 267–276.

Crooks, A., Pfoser, D., Jenkins, A., Croitoru, A., Stefanidis, A., Smith, D., . . . Lamprianidis, G. (2015). Crowdsourcing urban form and function. *International Journal of Geographical Information Science*, *29*(5), 720–741.

Culotta, A. (2010). Towards detecting influenza epidemics by analyzing Twitter messages. In *Proceedings of the First Workshop on Social Media Analytics*. Washington DC, USA: ACM.

Cummins, R. A., & Nistico, H. (2002). Maintaining life satisfaction: The role of positive cognitive bias. *Journal of Happiness Studies*, *3*, 37–69.

Czepkiewicz, M., Jankowski, P., & Młodkowski, M. (2017). Geo-questionnaires in urban planning: Recruitment methods, participant engagement, and data quality. *Cartography and Geographic Information Science*, *44*(6), 551–567.

Dalvi, N., Kumar, R., & Pang, B. (2012). Object matching in tweets with spatial models. In *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining* (pp. 43–53). Seattle, USA: ACM.

Dempsey, C., & Zimbalist, A. (2017). *Boston Olympics: How and why smart cities are passing on the torch* (1st ed.). Lebanon, USA: University Press of New England.

Diaz, F., Gamon, M., Hofman, J. M., Kiciman, E., & Rothschild, D. (2016). Online and social media data as an imperfect continuous panel survey. *PloS one*, *11*(1).

Dodds, P. S., Clark, E. M., Desu, S., Frank, M. R., Reagan, A. J., Williams, J. R., . . . & Danforth, C. M. (2015). Human language reveals a universal positivity bias. *Proceedings of the National Academy of Sciences*, *112*(8), 2389–2394.

Eisenstein, J. (2013). What to do about bad language on the internet. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human language technologies* (pp. 359–369). Atlanta, USA: The Association for Computational Linguistics.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2000). *Quantitative geography: Perspectives on spatial data analysis* (1st ed.). London, UK: Sage.

Frank, M. R., Mitchell, L., Dodds, P. S., & Danforth, C. M. (2013). Happiness and the patterns of life: A study of geolocated tweets. *Scientific Reports*, *3*(1), 2625.

Fraustino, J. D., Liu, B., & Jin, Y. (2012). *Social media use during disasters: A review of the knowledge base and gaps*. National Consortium for the Study of Terrorism and Responses to Terrorism, University of Maryland.

Fujisaka, T., Lee, R., & Sumiya, K. (2010). Exploring urban characteristics using movement history of mass mobile microbloggers. In *Proceedings of the Eleventh Workshop on Mobile Computing Systems & Applications* (pp. 13–18). Annapolis, USA: ACM.

Gayo-Avello, D. (2013). A meta-analysis of state-of-the-art electoral prediction from Twitter data. *Social Science Computer Review*, *31*(6), 649–679.

Gensim. (2017). *Gensim Python package*. Retrieved from https://radimrehurek.com/gensim

Getz, D., & Page, S. J. (2016). *Event studies: Theory, research and policy for planned events* (3rd ed). New York, USA: Routledge.

Girardin, F., Vaccari, A., Gerber, A., Biderman, A., & Ratti, C. (2009). Quantifying urban attractiveness from the distribution and density of digital footprints. *International Journal of Spatial Data Infrastructures Research*, *4*, 175–200.

Greater London Authority's DataStore. (2017). Statistical GIS boundary files for London. *London Datastore*. Retrieved from https://data.london.gov.uk/dataset/statistical-gis-boundary-files-london

Gupta, A., & Kumaraguru, P. (2012). Credibility ranking of tweets during high impact events. in *Proceedings of the 1st Workshop on Privacy and Security in Online Social Media*. (Article No. 2) Lyon, France: ACM.

Habidatum. (2017). Sport events & local economy.

Habidatum. Retrieved from https://habidatum.com/projects/sport-events-and-local-economy

Hall, P. (1966). *The world cities* (2nd ed.). London, UK: Weidenfeld and Nicolson.

Harvey, D. (1987). Flexible accumulation through urbanization: Reflections on "postmodernism" in the American city. *Antipode*, *19*(3), 260–286.

Hasan, S., & Ukkusuri, S. V. (2014). Urban activity pattern classification using topic models from online geo-location data. *Transportation Research Part C: Emerging Technologies*, *44*, 363–381.

Hawelka, B., Sitko, I., Beinat, E., Sobolevsky, S., Kazakopoulos, P., & Ratti, C. (2014). Geo-located Twitter as proxy for global mobility patterns. *Cartography and Geographic Information Science*, *41*(3), 260–271.

Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 168–177). Seattle, USA: ACM.

Kaufmann, B. (2015). If you want Olympics, do not ask the people. *Swissinfo.ch*. Retrieved from https://www.swissinfo.ch/eng/directdemocracy/democracy-and-olympic-spirit_if-you-want-the-olympics–do-not-ask-the-people/41595460

Keßler, C., Maué, P., Heuer, J. T., & Bartoschek, T. (2009). Bottom-up gazetteers: Learning from the implicit semantics of geotags. In K. Janowicz, M. Raubal, & S. Levashkin (Eds.), *GeoSpatial Semantics. GeoS 2009. Lecture Notes in Computer Science* (Vol. 5892). Berlin and Heidelberg: Springer.

Kim, J. W., Kim, D., Keegan, B., Kim, J. H., Kim, S., & Oh, A. (2015). Social media dynamics of global co-presence during the 2014 FIFA World Cup. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 2623–2632). Seoul: ACM.

Kinsella, S., Murdock, V., & Hare, N. O. (2011). "I'm Eating a Sandwich in Glasgow": Modeling locations with Tweets. In *Proceedings of the 3rd International Workshop on Search and Mining User-Generated Contents* (pp. 61–68). Glasgow, UK: ACM.

Lee, J. H., Gao, S., & Goulias, K. G. (2016). Can Twitter data be used to validate travel demand models? In *95th Annual Transportation Research Board Meeting* (pp. 1–27). Washington, DC, USA: TRB.

Lee, R., & Sumiya, K. (2010). Measuring geographical regularities of crowd behaviors for Twitter-based geo-social event detection. In *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location-Based Social Networks* (pp. 1–10). San Jose, USA: ACM.

Li, R., Lei, K. H., Khadiwala, R., & Chang, K. C.-C. (2012). TEDAS: A Twitter-based event detection and analysis system. In *Proceedings of the 28th International Conference on Data Engineering* (pp. 1273–1276). Washington, DC, USA: IEEE.

Li, L., Goodchild, M. F., & Xu, B. (2013). Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. *Cartography and Geographic Information Science*, *40*(2), 61–77.

Li, W., Serdyukov, P., de Vries, A. P., Eickhoff, C., & Larson, M. (2011). The where in the tweet. *Proceedings of the 20th ACM International Conference on Information and Knowledge Management* (pp. 2473–2476). Glasgow, UK: ACM.

Longley, P. A., Adnan, M., & Lansley, G. (2015). The geotemporal demographics of Twitter usage. *Environment and Planning A*, *47*(2), 465–484.

Malfas, M., Houlihan, B., & Theodoraki, E. (2004). Impacts of the Olympic Games as mega-events. *Proceedings of the Institution of Civil Engineers—Municipal Engineer*, *157*(3), 209–220.

Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J., & McClosky, D. (2014). The Stanford CoreNLP Natural Language Processing Toolkit. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System demonstrations* (pp. 55–60). Baltimore, USA: Association for Computational Linguistics.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, *50*(4), 370–396.

McGill. (2017). About urban planning. *McGill, School of Urban Planning*. Retrieved from https://mcgill.ca/urbanplanning/planning

Mellon, J., & Prosser, C. (2017). Twitter and Facebook are not representative of the general population: Political attitudes and demographics of British social media users. *Research & Politics*, *4*(3), 1–9.

Mitchell, L., Frank, M. R., Harris, K. D., Dodds, P. S., & Danforth, C. M. (2013). The geography of happiness: Connecting Twitter sentiment and expression, demographics, and objective characteristics of place. *PLoS ONE*, *8*(5).

Moore, J. (2015). When Denver rejected the Olympics in favour of the environment and economics. *The Guardian*. Retrieved from https://www.theguardian.com/sport/blog/2015/apr/07/when-denver-rejected-the-olympics-in-favour-of-the-environment-and-economics

Morstatter, F., Pfeffer, J., Liu, H., & Carley, K. M. (2013). Is the sample good enough? Comparing data from Twitter's streaming API with Twitter's firehose. In *Proceedings of the 7th International Conference on Weblogs and Social Media* (pp. 400–408). Cambridge, USA: AAAI Press.

Noulas, A., Scellato, S., Mascolo, C., & Pontil, M. (2011). Exploiting semantic annotations for clustering geographic areas and users in location-based social networks. In *Proceedings of the Social Mobile Web* (pp. 32–35). Barcelona, Spain: AAAI Press.

Pak, A., & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. In *Proceedings of the International Conference on Language Resources and Evaluation* (pp. 1320–1326). Valletta, Malta: European Language Resources Association.

Panteras, G., Wise, S., Lu, X., Croitoru, A., Crooks, A., &

Stefanidis, A. (2015). Triangulating social multimedia content for event localization using Flickr and Twitter. *Transactions in GIS*, *19*(5), 694–715.

Pennacchiotti, M., & Popescu, A. (2011). A machine learning approach to Twitter user classification. In *Proceedings of the Fifth International Conference on Weblogs and Social Media* (pp. 281–288). Barcelona, Spain: AAAI Press.

Porter, M. F. (1980). An algorithm for suffix stripping. *Program*, *14*(3), 130–137.

Preuss, H. (2007). The conceptualisation and measurement of mega sport event legacies. *Journal of Sport & Tourism*, *12*(3/4), 207–228.

Quercia, D., Ellis, J., Capra, L., & Crowcroft, J. (2012). Tracking "gross community happiness" from tweets. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (pp. 965–968). Seattle, USA: ACM.

Resch, B. (2013). People as sensors and collective sensing: Contextual observations complementing geosensor network measurements. In J. Krisp (Ed.), *Progress in location-based services* (pp. 391–406). Berlin and Heidelberg: Springer.

Resch, B., Summa, A., Zeile, P., & Strube, M. (2016). Citizen-centric urban planning through extracting emotion information from Twitter in an interdisciplinary space-time-linguistics algorithm. *Urban Planning*, *1*(2), 114–127.

Resch, B., Usländer, F., & Havas, C. (2017). Combining machine-learning topic models and spatiotemporal analysis of social media data for disaster footprint and damage assessment. *Cartography and Geographic Information Science*, 1–15.

Ritchie, B. W., Shipway, R., & Cleeve, B. (2009). Resident perceptions of mega-sporting events: A non-host city perspective of the 2012 London Olympic Games. *Journal of Sport & Tourism*, *14*(2/3), 143–167.

Ritterman, J., Osborne, M., & Klein, E. (2009). Using prediction markets and Twitter to predict a swine flu pandemic. In *Proceedings of the 1st International Workshop on Mining Social Media* (pp. 9–17).

Sagl, G., Resch, B., Hawelka, B., & Beinat, E. (2012). From social sensor data to collective human behaviour patterns–Analysing and visualising spatio-temporal dynamics in urban environments. In T. Jekel, A. Car, J. Strobl, & G. Griesebner (Eds.), *GI_Forum 2012: Geovizualisation, society and learning* (pp. 54–63). Berlin and Offenbach, Germany: Herbert Wichmann Verlag.

Sims, S. (2017). How Bostonians defeated the Olympics. *Citylab*. Retrieved from https://www.citylab.com/equity/2017/06/how-bostonians-defeated-the-olympics/529155

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

Steiger, E., Westerholt, R., Resch, B., & Zipf, A. (2015). Twitter as an indicator for whereabouts of people? Correlating Twitter with UK census data. *Computers, environment and urban systems*, *54*, 255–265.

Sui, D., & Goodchild, M. (2011). The convergence of GIS and social media: Challenges for GIScience. *International Journal of Geographical Information Science*, *25*(11), 1737–1748.

Sun, Y., Fan, H., Bakillah, M., & Zipf, A. (2015). Road-based travel recommendation using geo-tagged images. *Computers, Environment and Urban Systems*, *53*, 110–122.

Theodore, D. (2006). Sense of the city: An alternative approach to urbanism. *Journal of Architectural Education*, *60*(2), 69–70.

Twitter INC. (2017). Retrieved from https://developer.twitter.com/en/docs

Wang, H., Can, D., Kazemzadeh, A., Bar, F., & Narayanan, S. (2012). A system for real-time Twitter sentiment analysis of 2012 US presidential election cycle. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics* (pp. 115–120). Jeju, Republic of Korea: Association for Computational Linguistics.

Weng, J., & Lee, B.-S. (2011). Event detection in Twitter. In *Proceedings of the Fifth International Conference on Weblogs and Social Media* (pp. 401–408). Barcelona, Spain: AAAI Press.

Whitson, D., & Macintosh, D. (1993). Becoming a world-class city: Hallmark events and sport franchises in the growth strategies of western Canadian cities. *Sociology of Sport Journal*, *10*(3), 221–240.

Yu, Y., & Wang, X. (2015). World Cup 2014 in the Twitter world: A big data analysis of sentiments in US sports fans' tweets. *Computers in Human Behavior*, *48*, 392–400.

Zeile, P., Resch, B., Exner, J.-P., & Sagl, G. (2015). Urban emotions: Benefits and risks in using human sensory assessment for the extraction of contextual emotion information in urban planning. In S. Geertman, J. Ferreira,R. Goodspeed, & J. Stillwell (Eds.), *Planning support systems and smart cities* (pp. 209–225). Switzerland: Springer International Publishing.

Zhang, X., Fuehres, H., & Gloor, P. A. (2011). Predicting stock market indicators through Twitter "I hope it is not as bad as I fear". *Procedia: Social and Behavioral Sciences*, *26*, 55–62.

Zhang, Z., Ni, M., He, Q., & Gao, J. (2016). *Mining transportation information from social media for planned and unplanned events*. Buffalo, NY, USA: University at Buffalo.

Zhao, W. X., Jiang, J., Weng, J., He, J., Lim, E.-P., Yan, H., & Li, X. (2011). Comparing Twitter and traditional media using topic models. In *Proceedings of 33rd European Conference on IR Research* (pp. 338–349). Dublin, Ireland: Springer Verlag.

Zhao, S., Zhong, L., Wickramasuriya, J., & Vasudevan, V. (2011). *Human as real-time sensors of social and physical events: A case study of Twitter and sports games* (Technical Report TR0620-2011). Rice University and Motorola Labs.

About the Authors

Anna Kovács-Győri has been a PhD candidate at the University of Salzburg, Department of Geoinformatics–Z_GIS since 2015. She studied Geography with a specialization in GIS at the Department of Physical Geography and Geoinformatics, University of Szeged, Hungary. During her PhD studies, she investigates the quality of the person-environment relationship in the urban context. Her research focuses on livability assessment and the role of spatiality, mobility, human scale, and health in it to promote urban planning and decision-making.

Alina Ristea is a fully funded PhD student at the Doctoral College GIScience, and her research topic is "Integration and Evaluation of Social Media Data into Crime Prediction Models", under the supervision of Professor Michael Leitner. Her background studies are in Geography–Cartography-Cadastral Elements (Bachelor) and GIS (Master) from the University of Bucharest, Romania.

Clemens Havas is a PhD student with a background in computer science at University of Salzburg's Department of Geoinformatics–Z_GIS since 2017. His research interests lie in the fields of machine learning, social media and geospatial analysis. Besides doing his PhD programme, Clemens Havas has been working for various projects funded by international organizations/programmes (e.g., Horizon 2020) by using new data sources to assess the impact of disasters.

Bernd Resch is an Assistant Professor at University of Salzburg's Department of Geoinformatics—Z_GIS and a Visiting Fellow at Harvard University (USA). His research interests revolve around fusing data from human and technical sensors, including the analysis of social media. Amongst a variety of other functions, Bernd Resch is Editorial Board Member of the *International Journal of Health Geographics*, Associated Faculty Member of the doctoral college "GIScience", and Executive Board member of Spatial Services GmbH.

Pablo Cabrera-Barona is Assistant Professor at the Institute of Higher National Studies (IAEN), and visitor Professor at the Latin American Social Sciences Institute (FLACSO) and San Francisco de Quito University (USFQ). His main interests lie in understanding different interactions between geography, economics, ecology, health and development processes. He is used to working at different scales, considering different social contexts.

Supplementary Files

Hourly spatiotemporal patterns (KDE)—for residents and visitors, positive and negative sentiments

The first row in each figure shows morning patterns (8:00 a.m.), the second one represents one of the most active hours in the data (12:00 noon) while the third and fourth represent the early (6:00 p.m.) and late evening (11:00 p.m.). Red spots are the densest areas in the given hour and category, while yellow shows slightly lower density. Blue represents those areas where based on the KDE algorithm, the density in a cell is lower than the expected. The cell size was in each map 200 m^2. The tables provide general information on the cell values' statistics, and it can aid the interpretation of how high the differences were between blue and red areas. The higher the standard deviation is, the higher the difference in the density, while using the maximum and mean values, the different maps become comparable.

Table 1. Values for the hot spot analysis using KDE for residents' negative tweets for specific hours.

	before	OG	after		before	OG	after
8:00 a.m.	n = 366 mean = 0.11 stdev = 0.25	n = 287 mean = 0.10 stdev = 0.19	n = 372 mean = 0.10 stdev = 0.21	**6:00 p.m.**	n = 365 mean = 0.11 stdev = 0.23	n = 323 mean = 0.10 stdev = 0.19	n = 351 mean = 0.10 stdev = 0.16
12:00 noon	n = 351 mean = 0.11 stdev = 0.18	n = 400 mean = 0.12 stdev = 0.28	n = 304 mean = 0.08 stdev = 0.17	**11:00 p.m.**	n = 418 mean = 0.12 stdev = 0.21	n = 364 mean = 0.11 stdev = 0.16	n = 288 mean = 0.08 stdev = 0.13

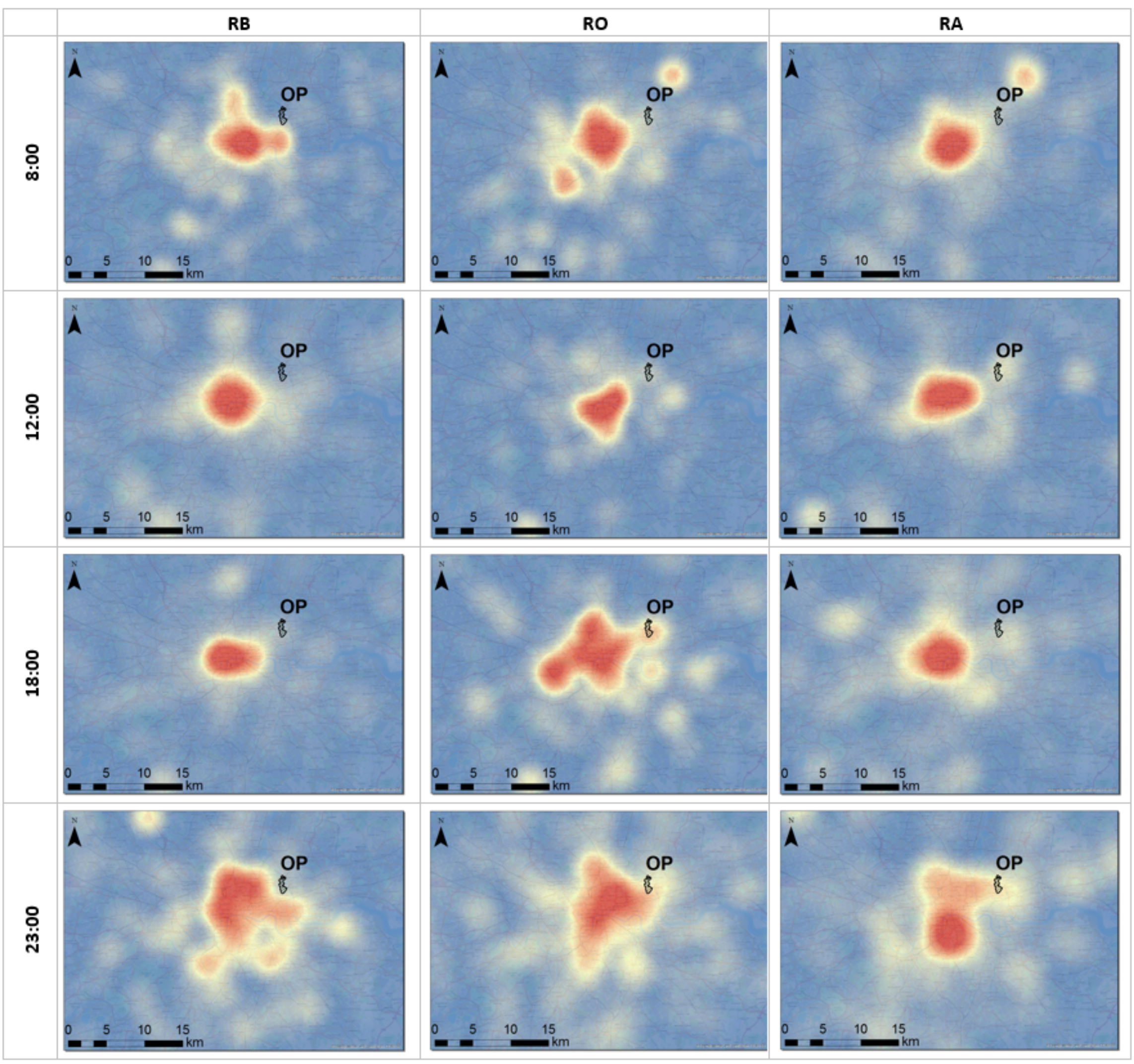


Figure 1. Hot spots of residents' negative tweets for specific hours.

Table 2. Values for the hot spot analysis using KDE for visitors' negative tweets for specific hours.

	before	OG	after		before	OG	after
8:00 a.m.	n = 217 mean = 0.06 stdev = 0.1	n = 171 mean = 0.05 stdev = 0.09	n = 317 mean = 0.09 stdev = 0.11	**6:00 p.m.**	n = 228 mean = 0.06 stdev = 0.13	n = 202 mean = 0.06 stdev = 0.11	n = 330 mean = 0.09 stdev = 0.15
12:00 noon	n = 236 mean = 0.07 stdev = 0.18	n = 260 mean = 0.08 stdev = 0.15	n = 369 mean = 0.11 stdev = 0.17	**11:00 p.m.**	n = 248 mean = 0.08 stdev = 0.21	n = 238 mean = 0.07 stdev = 0.11	n = 445 mean = 0.12 stdev = 0.16

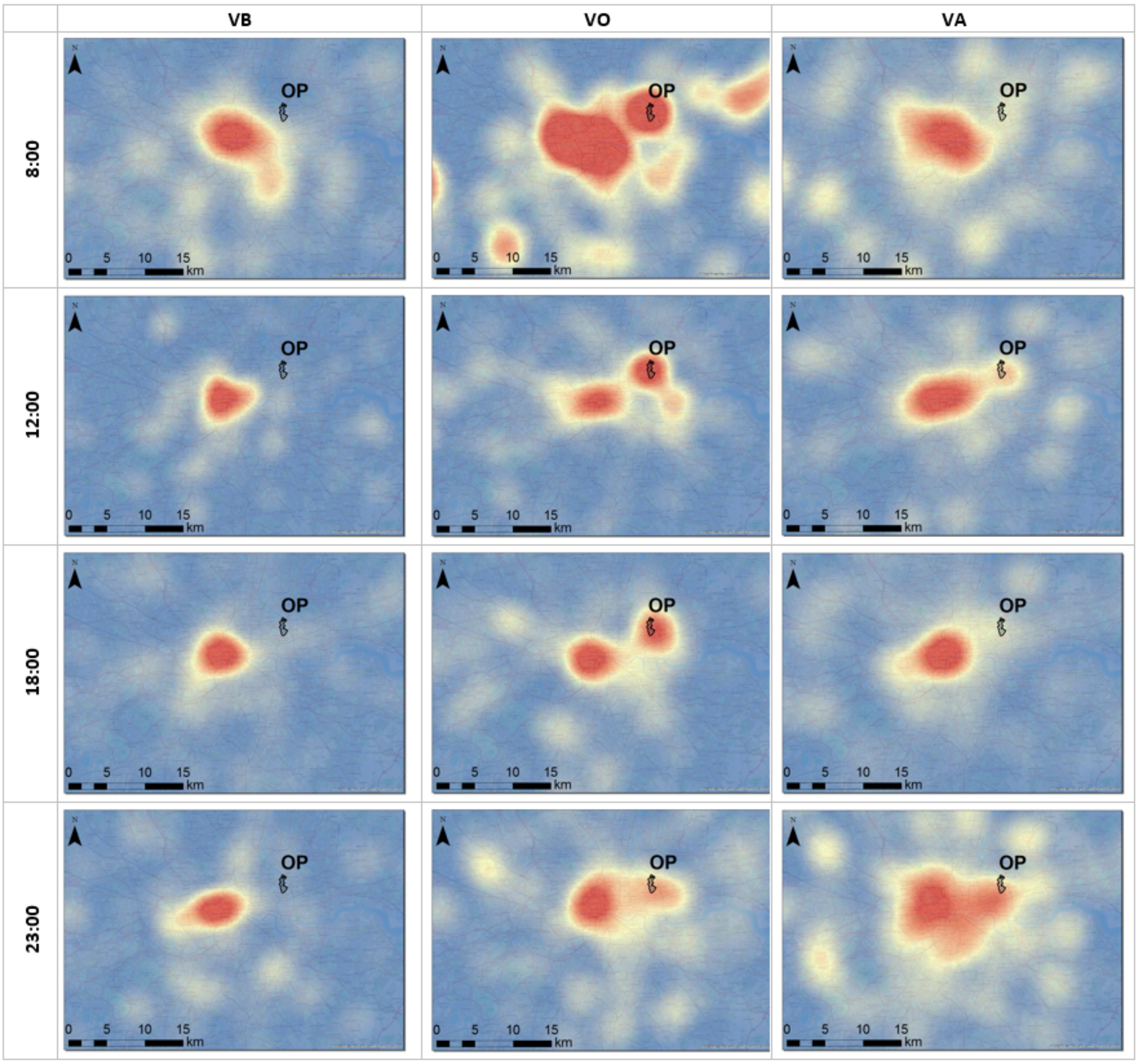


Figure 2. Hot spots of visitors' negative tweets for specific hours.

Table 3. Values for the hot spot analysis using KDE for residents' positive tweets for specific hours.

	before	OG	after		before	OG	after
8:00 a.m.	n = 839 mean = 0.25 stdev = 0.7	n = 819 mean = 0.24 stdev = 0.53	n = 809 mean = 0.23 stdev = 0.47	**6:00 p.m.**	n = 773 mean = 0.23 stdev = 0.62	n = 868 mean = 0.25 stdev = 0.62	n = 661 mean = 0.19 stdev = 0.47
12:00 noon	n = 819 mean = 0.24 stdev = 0.83	n = 939 mean = 0.27 stdev = 0.68	n = 727 mean = 0.23 stdev = 0.54	**11:00 p.m.**	n = 702 mean = 0.20 stdev = 0.41	n = 828 mean = 0.24 stdev = 0.38	n = 565 mean = 0.16 stdev = 0.27

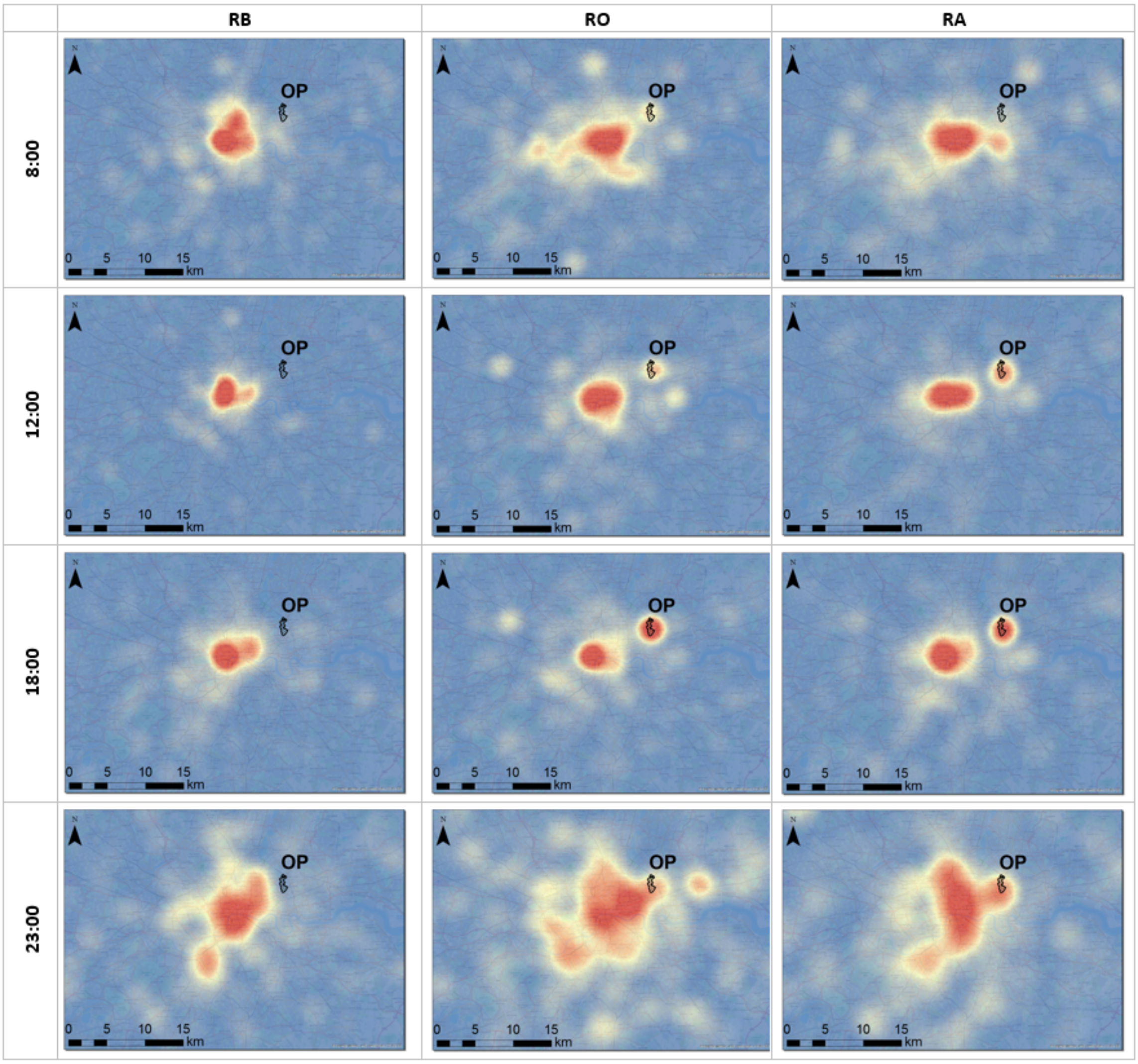


Figure 3. Hot spots of residents' positive tweets for specific hours.

Table 4. Values for the hot spot analysis using KDE for visitors' positive tweets for specific hours.

	before	OG	after		before	OG	after
8:00 a.m.	n = 452 mean = 0.13 stdev = 0.28	n = 498 mean = 0.15 stdev = 0.32	n = 753 mean = 0.21 stdev = 0.36	**6:00 p.m.**	n = 483 mean = 0.14 stdev = 0.46	n = 562 mean = 0.16 stdev = 0.56	n = 684 mean = 0.20 stdev = 0.48
12:00 noon	n = 565 mean = 0.16 stdev = 0.48	n = 821 mean = 0.24 stdev = 0.78	n = 737 mean = 0.21 stdev = 0.55	**11:00 p.m.**	n = 523 mean = 0.15 stdev = 0.38	n = 577 mean = 0.16 stdev = 0.36	n = 706 mean = 0.20 stdev = 0.34

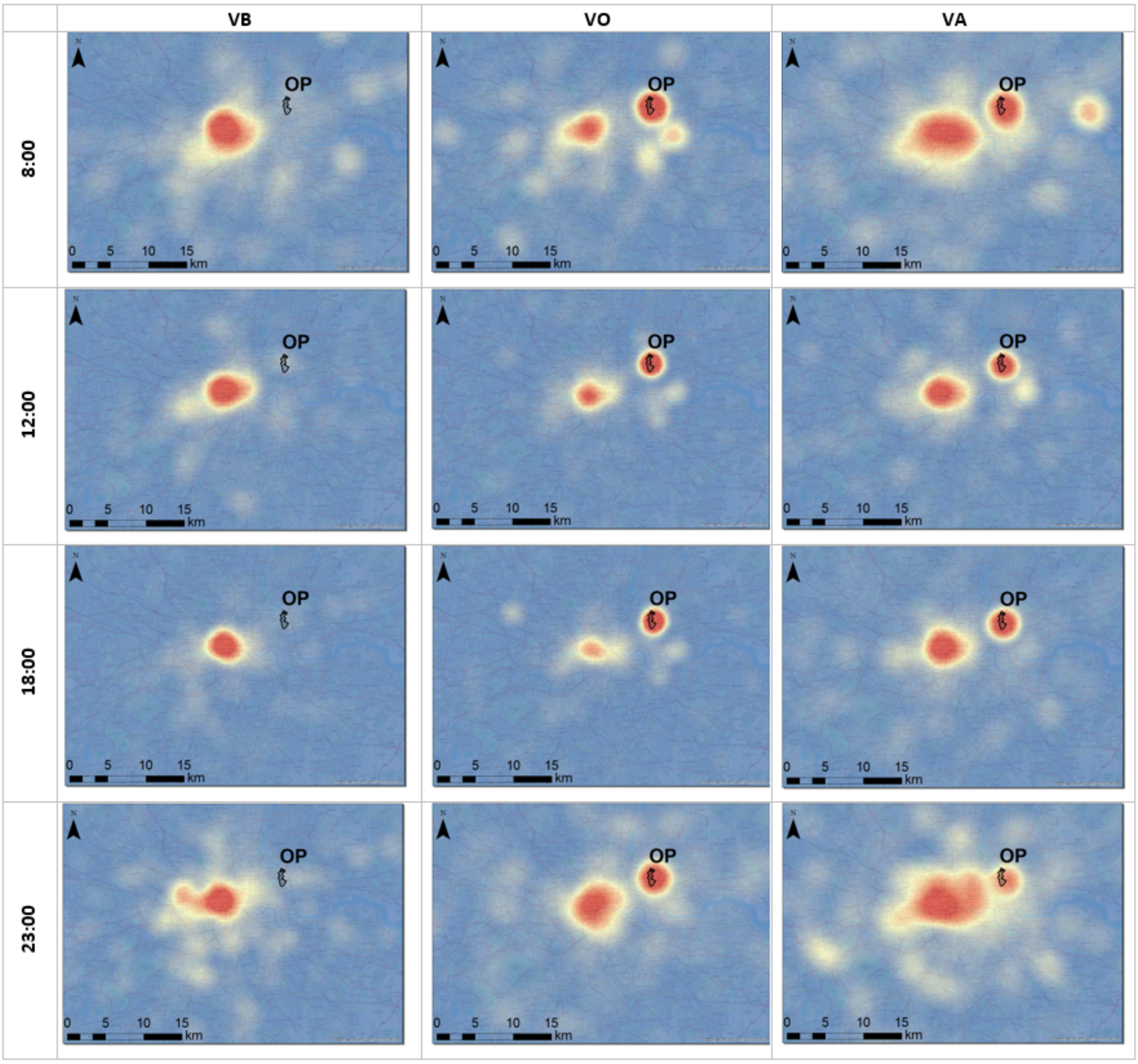


Figure 4. Hot spots of visitors' positive tweets for specific hours.

Urban Planning (ISSN: 2183–7635)
2018, Volume 3, Issue 1, Pages 100–115
DOI: 10.17645/up.v3i1.1266

Article

Crowdsourcing Local Knowledge with PPGIS and Social Media for Urban Planning to Reveal Intangible Cultural Heritage

Pilvi Nummi

Department of Built Environment, Aalto University, 02150 Espoo, Finland; E-Mail: pilvi.nummi@aalto.fi

Submitted: 15 November 2017 | Accepted: 13 March 2018 | Published: 29 March 2018

Abstract
In participatory urban planning, understanding local stakeholders' viewpoints is central, and, thus, gathering local knowledge has become a frequent task in planning practice. However, the built cultural heritage is usually evaluated by experts neglecting the values and opinions of citizens. In this study, a crowdsourcing model for assessing local residents' viewpoints and values related to the built cultural heritage of Nikkilä was developed. The aim was to find out if crowdsourcing with public participation GIS and social media is a functional method for revealing local people's values, place-based memories and experiences. In the case study, non-professional knowledge was compared with expert knowledge and valuable knowledge about the intangible aspects of the built cultural heritage was reached through place-based memories. Apart from that, social media provided visual representations of place-based experiences and a tool for building a collective memory. Based on the results, it is evident that a multi-method crowdsourcing model can be a functional model for crowdsourcing local knowledge. However, there are several challenges in analysing data and using the knowledge in urban planning.

Keywords
crowdsourcing; cultural heritage; place-based memories; PPGIS; social media; urban planning

Issue
This article is part of the issue "Crowdsourced Data and Social Media for Participatory Urban Planning", edited by Bernd Resch (University of Salzburg, Austria), Peter Zeile (Karlsruhe Institute of Technology, Germany) and Ourania Kounadi (University of Salzburg, Austria).

1. Introduction

Crowdsourcing can be a powerful tool for enhancing public participation in urban planning processes (Brabham, 2009). In general, crowdsourcing can be considered as an activity of using the power of the crowd to accomplish a task or to solve a problem (Nakatsu, Grossman, & Charamblos, 2014). The concept was first introduced by Jeff Howe in Wired magazine in 2006, when he wrote about businesses outsourcing tasks to undefined groups of people (Howe, 2006).

In academic research, there are various interpretations of crowdsourcing. Recently, the concept of crowdsourcing has been used in research related to public participation GIS (PPGIS) in the meaning of gathering data with map surveys (e.g., López-Aparicio, Vogt, Schneider, Kahila-Tani, & Broberg, 2017; Pánek & Benediktsson, 2017) and extracting knowledge from user-generated social media data (e.g., Dunkel, 2015; Zhou & Zhang, 2016). Contrary to this, Brabham (2009) and Seltzer and Mahmoudi (2012) argue that collecting data for planning with web-based surveys or using user-generated social media data should not be considered as crowdsourcing. They emphasise that crowdsourcing is about problem-solving and, in urban planning, crowdsourcing should include planning solutions to answer planning problems. However, based on a wider definition of crowdsourcing presented by Nakatsu et al. (2014), any task related to the job (i.e., planning) can be crowdsourced.

Nevertheless, some researchers have argued that more empirical research in close relation to real planning situations is needed to prove if crowdsourcing methods really are applicable and useful in urban planning and decision making (Nielek, Ciastek, & Kope, 2017).

This study fills that gap by applying a crowdsourcing method in an actual urban planning case study in Nikkilä, Sipoo, Finland.

In this study, knowledge gathering with PPGIS and social media is considered as crowdsourcing. In the case study Nikkilä Memories, a crowdsourcing model for collecting local knowledge related to people's memories, experiences and values is developed and tested in Nikkilä, Sipoo. Social media and PPGIS are used for collecting local people's place-based memories related to old buildings and places of Nikkilä. In addition to the collected data, the author uses her own experiences from urban planning practices as a source for this study.

1.1. Introduction to the Case Study Nikkilä Memories

Nikkilä is the administrative centre of Sipoo, a growing municipality within the metropolitan area of Helsinki, Finland. In Sipoo, there are currently 20,000 inhabitants, of which approximately 4,400 live in Nikkilä. Despite its small population size and rural character, Sipoo is at the moment one of the fastest growing municipalities in Finland with 2% annual population growth rate. The centre of Nikkilä is undergoing a transformation: the population rate will be doubled in the coming decades, and, thus, the existing structure of the centre will be densified and new housing areas will be built close to the centre. Currently, an outline plan of the whole area is in the making, to define the areas for densification and enlargement of the town. The Nikkilä Memories case study provides local knowledge, especially for the outline planning process, but also for future detailed planning of Nikkilä.

In the post-industrial cities, cultural aspects such as cultural heritage have been identified as an important resource for urban development and planning. Built cultural heritage is an evident asset when cities are aiming to create and maintain a distinctive and authentic sense of place (Bianchini & Ghilardi, 2007). The cultural resources of Nikkilä have been studied recently (Nummi & Tzoulas, 2015), and, based on that, built cultural heritage has been identified as a cornerstone of the identity of the place. Therefore, it is relevant to further study the values and meanings related to the cultural heritage of Nikkilä.

Furthermore, it is relevant to study and understand intangible aspects of cultural heritage. Zukin (2012) argues that 'advocates working within the framework of cities' historic preservation laws generally focus on protecting the tangible heritage of individual buildings and districts' while the intangible heritage is neglected. Historic monuments and buildings are being preserved, whereas urban places that do not possess tangible historical values are underappreciated. Hence, to reveal the intangible aspects of the cultural heritage of Nikkilä, it is relevant to understand the local culture and local stakeholders' viewpoints.

In the case study of Nikkilä, crowdsourcing is used to collect and share place-based memories and experiences to find out what the value of the built cultural heritage is and what elements form the intangible cultural heritage of Nikkilä. The idea of collecting place-based memories came together in a discussion between the author and the planners of Sipoo. The planners pointed out that there was a need to update and complete the expert inventory of the built cultural heritage as the data was inadequate for assessing the significance of old buildings. They also wanted to engage local people in the evaluation of the buildings, and, together with the author, they came up with an idea of crowdsourcing.

The novelty of this research is that it combines two complementary methods and data sources, i.e., social media and PPGIS data, in the crowdsourcing model. Apart from that, in relation to the cultural heritage, this study brings together experts and non-professional knowledge as parallel and comparable.

2. Background

2.1. Participatory Urban Planning

In the urban planning context, crowdsourcing can be seen as a model for enhancing public participation (Brabham, 2009). Planning theorists, like Patsy Healey (1997, 1999), define participatory planning as a communicative and human-centred approach to urban planning and decision making. Participatory planning is connected to deliberative democracy, a model of democracy that emphasises public debate and discussion as a basis for fair decision making. The framework of collaborative planning, presented by Healey (1997, 1999), is a communicative model for participatory urban planning. Healey describes planning as 'an interactive process, involving communicative work among participants, during which issues, problems, strategies and policy ideas are given form and meaning' (Healey, 1997, p. 91). She also points out that knowledge building that aims at shared understanding is a central part of a collaborative planning process (Healey, 1999). Furthermore, Healey (1997) emphasises acknowledging and accepting different types of knowledge, both expert and non-expert, into the planning process. Van Herzele (2004) describes local knowledge as non-professional knowledge expressed by participants (particularly local residents) in a planning process. Furthermore, Kahila and Kyttä (2008) argue that experiential knowledge generated by local people can be considered as local knowledge as well. Based on this, local place-based memories and experiences produced in the case study Nikkilä Memories can be seen as local knowledge.

The adoption of participatory urban planning practices in Finland has been advanced by the land use and building law that came into effect in 2000. It is required that everyone in Finland should have the opportunity to participate in the planning processes that are affecting their everyday lives or living environment (Finlex, 1999). It is mandatory to present plan proposals publicly and provide citizens with a possibility to give feedback about the proposed planning solution, but it is not obligatory, for example, to engage citizens in goal setting in the beginning of the planning process or in designing planning solutions during the planning process. Despite the good intentions of the act, it has produced planning practices where participation takes place at a very late stage of planning, and the citizens' role remains narrow, an opinion giver. It is widely acknowledged that successful results require participation in an early phase of planning (Eräranta & Staffans, 2015). In the case study of Nikkilä Memories, the crowdsourcing method is tested in an early phase of planning to answer this practical question.

2.2. Crowdsourcing in Urban Planning Context

The principal definition of crowdsourcing in the urban planning context is problem-solving. Both Brabham (2009) and Seltzer and Mahmoudi (2012) present models of crowdsourcing solutions for planning problems. Seltzer and Mahmoudi do not include social media, PPGIS or other survey methods in crowdsourcing. They argue that crowdsourcing should always contain ideas generation and solution selection phases, which usually do not exist in processes that focus on gathering local knowledge. Brabham (2009) describes crowdsourcing in urban planning as a tool for outsourcing the design of planning solutions. Furthermore, he argues that it is a question of empowering citizens by letting them also vet the planning solutions created by the crowd. It is worth mentioning that, according to Brabham, crowdsourcing is not a bottom-up approach to urban planning. On the contrary, it is very much a top-down steered process with a well-defined and focused task that the crowd is participating in.

In recent crowdsourcing studies, examples of local knowledge gathering and methods supporting non-experts to design planning proposals are presented. For example, Mueller, Lu, Chirkin, Klein and Schmitt (2018) present a model and toolkit (Quick Urban Analysis Kit) for crowdsourcing citizen feedback, ideas and wishes. They developed a participatory planning strategy which combined a co-design approach with crowdsourcing methods and introduced a design tool that enables non-experts to do simple design tasks. Examples of knowledge gathering for urban planning relate to various subjects. For example, Yang and Ng (2017) have tested the potential of using crowdsourced user-generated data to monitor urban rainfall. They found that crowdsourced data lead to a more accurate modelling of storm water flows as compared to rain gauge data. Pánek and Benediktsson (2017) applied an 'emotional mapping' methodology to study cyclists' opinions about cycling routes and places that they encounter in Reykjavík, Iceland. They describe their data collection method as geographical crowdsourcing, and argue that this kind of participatory planning support system can help to bridge the gap between planners and citizens. López-Aparicio et al. (2017) used a PPGIS tool for collecting information about wood burning for residential heating in Norway, and described the method as crowdsourcing as well.

The development and research of PPGIS methods, web map tools for gathering local knowledge, have been going on for a decade (Kahila-Tani, Broberg, Kyttä, & Tyger, 2016). Brown (2015) argues that crowdsourcing with PPGIS tools is starting to become more frequent in urban planning practices. Crowdsourcing with PPGIS tools is closely related with the concept of volunteered geographical information (VGI), which means the process of users voluntarily producing geographic data, such as marking geographic features or objects (e.g., updating OpenStreetMap) or adding geodata to objects shared in social media (Elwood, Goodchild, & Sui, 2012). Taking into account Brabham's (2009) definition of crowdsourcing, VGI can also be considered as crowdsourcing if it is steered top-down, for example, by the municipal authorities in the case of urban planning.

On the grounds of previous definitions of crowdsourcing and examples, social media should not be considered as crowdsourcing, because the users are not accomplishing any predefined tasks or trying to solve a problem. However, in many studies that use user-generated social media data as a data source for an analysis, this is presented as a form of crowdsourcing, as will be reported in Section 2.3. Also, a social media platform can be used for crowdsourcing if the users are contributing to a task, for example, taking part in data gathering by posting pictures on Instagram, as in the case study of Nikkilä Memories.

2.3. Social Media Data Analysis in Urban Planning

Using social media in knowledge gathering for urban planning is relatively new and there are few studies related to actual planning processes using social media data (Nummi, 2017). Generally, in public administration, social media tools are used for information dissemination and collaboration, and Khan (2015) argues that models for social media-based government are found useful for increasing public participation. Khan presents different relationships between citizens and government, and points out that citizens can, for example, act as an informational source via social media channels by providing feedback or expert opinion for government.

The use of social media has been increasing over the last decade and new services are being implemented. Social media have become a part of many people's everyday life: for example, in Finland, more than 50% of inhabitants have a profile in a social media channel (Official Statistics of Finland, 2017). For urban planning, this offers opportunities to communicate with citizens and to study various aspects of citizens' experiences, opinions and feelings. In other words, social media can be used as an interaction tool or a data source in urban planning.

When social media are seen as a data source, as in this study, it is relevant to contemplate social media data analysis methods (SMDA). These are computational methods for analysing user-generated content from social media (Nummi, 2017). Recently, in academia, there has been an increasing interest in adopting social media analysis methods for urban planning. For example, Tasse and Hong (2014) presented opportunities to use geotagged social media data to better understand cities, and a variety of SMDA methods have been studied in case studies in different parts of the world. The increasing numbers of research papers in this field suggest the topicality of this subject. These methods can be categorised in four themes: opinion mining, place-based experiences, understanding people's behaviour in a city, and the city structure based on social media data (Nummi, 2017).

For example, people's place-based experiences have been studied by Dunkel (2015), who developed a GIS method for analysing people's landscape experiences and values using crowdsourced data from Instagram. Apart from that, urban sounds and smells (Aiello, Schifanella, Quercia, & Aletta, 2016; Quercia, Aiello, & Schifanella, 2016), feeling of stress (Sykora et al., 2015) and emotions (Resch, Summa, Zeile, & Strube, 2016) have been studied to help planners to understand how people experience the city. Opinion mining has been used by Campagna and colleagues (Campagna, 2014; Campagna, Floris, Massa, Girsheva, & Ivanov, 2015), who created a special social media platform to listen to citizens' viewpoints in relation to city planning in Cagliari, Italy, and by López-Ornelas and Zaragoza (2015), who studied opinions and feelings related to a new airport in Mexico City. Points of interest (POI) are an obvious source of planning-related data that describes crowd behaviour in a city and, thus, provide insight into places and areas that attract people (Hu et al., 2015; Paldino, Bojic, Sobolevsky, Ratti, & González, 2015).

Analysis methods that reveal the structure of the city are often based on location-based (e.g., geotagged) social media data. For example, Frias-Martinez and Frias-Martinez (2014) studied geotagged Tweets to find out the actual land use of areas based on people's activities. They were able to analyse areas that are used for business, leisure, housing and nightlife with a GIS analysis. Methods that detect similarities between areas are also enlightening the city structure from users' point of view. For example, Cranshaw, Schwartz, Hong and Sadeh (2012) developed a tool called Livehoods that visualises on a map areas that are similar based on residents' activities.

3. Research Questions and Methods

The overall research problem is what kind of crowdsourcing model can reveal local people's values and place-based memories related to the built cultural heritage. The need for this case study came from the urban planners of Sipoo, as they had pointed out the need to better understand the value and meaning of old buildings and places in Nikkilä for local inhabitants. An inventory of cultural heritage objects and areas was conducted earlier, in 2005, by experts in built cultural heritage and architecture (Municipality of Sipoo, 2006). The weakness of such expert examination is that it does not take residents' viewpoints into account. One important question in the case study was whether the values of residents are in line with this expert evaluation or not. In participatory urban planning, identifying local people's viewpoints and values is central to be able to focus the development goals and draw alignments in a culturally and socially sustainable way.

The research questions were posed from both research and planning perspectives. The actual research questions were related to crowdsourcing model and methods:

- What kind of crowdsourcing model can be used for gathering local place-based memories, values and experiences related to the built cultural heritage?
- Is crowdsourcing a functional model for this task?
- What is the value of using social media in crowdsourcing?
- What kind of challenges and benefits does this crowdsourcing model produce?

From the urban planning perspective, the following questions are studied:

- Which old buildings are valuable to local people in Nikkilä?
- Are experts' evaluations and the values of non-professional people in line or do they differ?
- What does the data reveal about the intangible aspects of cultural heritage?
- How can the results from the crowdsourcing be reflected in planning?

In the case study of Nikkilä Memories, these questions are approached with a multi-method approach that combines a map questionnaire and social media data. The map questionnaire was created with the Maptionnaire tool. Social media were used to facilitate people to share their memories from Nikkilä with hashtag #muistojennikkilä, and a local Facebook group related to old buildings in Sipoo was observed.

The aim of this case study was to examine how local people value cultural heritage objects (e.g., historical buildings and cultural landscape areas in Nikkilä) and what kind of memories they have related to the built cultural heritage. The idea behind the approach was that people's memories are related to the appreciation they feel towards the cultural heritage, and that memories can reveal the intangible cultural heritage of Nikkilä.

4. The Case Study Nikkilä Memories

In the following sections, the starting points, methods and data of the case study Nikkilä Memories are presented. First, the village of Nikkilä is described, and then crowdsourcing methods (map questionnaire, social media monitoring and observation of a self-organized Facebook group) and the analysis method are presented.

4.1. Nikkilä—A Small Town in the Centre of Sipoo, Finland

Nikkilä is a town with currently only approximately 4,400 inhabitants. The area is undergoing major transformation, because the aim of the municipality is to more than double the population within the next 10 years. This, of course, means that the environment will change significantly, and, thus, it is important to engage local people in the planning of the area. The case study of Nikkilä Memories is part of an ongoing participatory process started a few years ago (Nummi & Tzoulas, 2015).

The areal image of Nikkilä (Figure 1) describes the physical characteristics of the village. The dense centre of Nikkilä is surrounded by rural agricultural areas, fields and forests. In the foreground of the picture, a new housing area of Sipoon Jokilaakso (the River Valley of Sipoo) is under construction. Apart from these expansion areas, the existing structure of Nikkilä is being densified. Old department and office buildings from the 1980s have been pulled down and replaced by new apartment buildings in the centre of Nikkilä (Figure 2, above).

The history of Nikkilä derives from the mediaeval age. An old mediaeval stone church (Figure 2) built in the fifteenth century is the oldest building in Nikkilä. The built cultural heritage is diverse; along the main street, Iso Kylätie, there are several old wooden buildings from the late nineteenth or early twentieth century (Figure 2). In the twentieth century, development of the village was strongly affected by the establishment of a mental hospital in 1914. The hospital operated in Nikkilä until 1999, and, after it was closed, the hospital area was renovated to a picturesque housing area, but still the termination of the hospital decreased the population and liveability of Nikkilä.

4.2. Crowdsourcing Model

The elementary idea of the crowdsourcing project was to set the expert-driven cultural heritage inventory under public evaluation with a map questionnaire. Apart from that, people's place-based memories and experiences were crowdsourced. A map questionnaire was developed with a PPGIS tool called Maptionnaire, and it was used in parallel with social media to offer multiple ways for people to express their thoughts. In social media, people were encouraged to share their memories from Nikkilä with the hashtag #muistojennikkilä. Both of these methods are explained in detail in the next sections.

4.3. Map Questionnaire

The map questionnaire was open from 21 March–31 July 2016 and 186 answers were received. The data consisted of more than 700 evaluations of old buildings, 39 written memories, 12 images and 106 markings for personally important buildings. The questionnaire was an open web questionnaire and, thus, the data is biased towards those participants who are most interested in the local built cultural heritage of Nikkilä. This kind of methodology can be considered as a crowdsourcing data collection (not as a survey method) as the sample of respondents is not representative.

Figure 1. Areal photo of Nikkilä. Copyright: Municipality of Sipoo, Ilmakuva Vallas Oy/Hannu Vallas. Used with permission.

Figure 2. Old and new buildings in Nikkilä. Above: New apartment buildings in the heart of Nikkilä. Copyright: Suvi Suovaara and Municipality of Sipoo; below: the old mediaeval stone church from the fifteenth century (picture by the author) and old wooden house in the main street of Nikkilä (Iso Kylätie), Copyright: Municipality of Sipoo. Used with permission.

The questionnaire was promoted in various ways: first, the preliminary version of the questionnaire was tested in an event for local residents of Nikkilä. Then, the project was launched and local media were informed, and the URL address of the questionnaire was shared in social media (e.g., the Facebook page of the municipality). Also, Facebook marketing targeted for users living in Sipoo was used to reach the respondents.

Table 1 shows the overview of respondents of the questionnaire. Surprisingly, although the map questionnaire was promoted in Sipoo, altogether 38% of the respondents came from outside the municipality, and only 32% were local residents of Sipoo. It is evident that many former residents answered the questionnaire. The age of the respondents centres around age groups from 26 to 65. Younger and older respondents were in the minority. Women were more active in responding (62%) than men (38%).

The map questionnaire comprised four parts: 1) evaluating old buildings (built cultural heritage) as shown in Figure 3; 2) adding personally important buildings on the map; 3) evaluating culturally important landscape areas; and 4) sharing memories and stories about Nikkilä. This article focuses on the built cultural heritage (parts 1 and 2) and people's place-based memories (part 4).

Table 1. Overview of the PPGIS respondents.

	Percentage (%)
Gender	
Male	38%
Female	62%
Age	
under 18	1%
19–25	9%
26–35	18%
36–45	20%
46–55	26%
56–65	18%
66–75	6%
over 75	1%
Place of residence	
Nikkilä	32%
Other place in Sipoo	30%
Somewhere else	38%

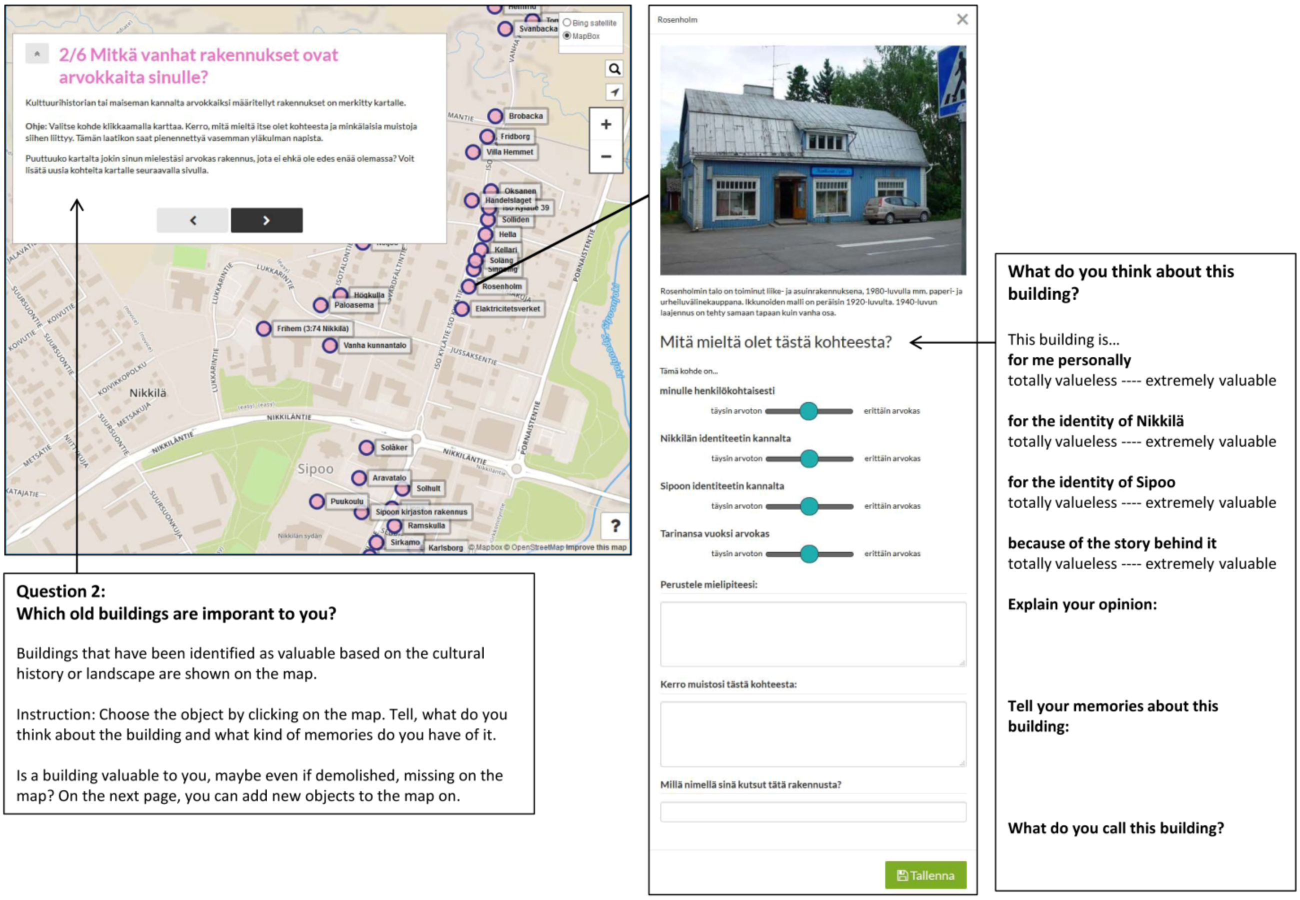


Figure 3. A screen of Nikkilä Memories map questionnaire. Copyright: Municipality of Sipoo and Mapita Oy. Used with permission.

The screenshot (Figure 3) represents the part of the questionnaire where the respondent was able to evaluate the cultural heritage buildings and to share memories about the building. The data was used to assess how people's evaluations differ from the experts' evaluation of the built heritage. Apart from that, memories linked to the buildings were analysed to understand the meaning of the buildings.

4.4. Social Media Monitoring

Social media were used as a crowdsourcing tool by encouraging people to share their memories from Nikkilä with hashtag #muistojennikkilä on Instagram, Twitter and Facebook. An Instagram profile called Nikkilä Memories was opened in April 2016 to facilitate crowdsourcing. With that profile, the author forwarded memories collected with the map questionnaire to inspire people to share their memories. ViidakkoMonitor, a social media monitoring tool developed by a Finnish IT company, Koodiviidakko Oy, was used for following the hashtag #muistojennikkilä in all the channels from June 2016 to June 2017.

During the time from June 2016 to June 2017, people shared in total 228 posts with the hashtag. Posts were manually classified by their content, and, as a result, 28% of the posts were related to landscapes, 22% to buildings, 19% to events organized in Nikkilä and 12% to art or culture.

Instagram was the primary tool, with 191 posts, and it appeared to complement the PPGIS data in an essential way by providing visual representations of peoples' place-based experiences. The major flaw with the social media data was that it did not contain any geodata. Apart from that, it is evident that only a small group of active residents were posting content.

4.5. Observation of a Local Facebook Group

A local Facebook group called 'Old buildings in Sipoo' (*Sipoon vanhat rakennukset–Gamla byggnader i Sibbo*) was observed to find out if people shared their memories or discussed the built cultural heritage on Facebook. The observation was executed manually during a two-month period in spring 2016.

'Old buildings in Sipoo' is a closed Facebook group with more than 800 members. In Facebook, anyone can find closed groups and see their members but only members can see posts in that group (Facebook, 2017). Closed or secret groups cannot be crawled with social media monitors and, to protect the privacy of users, Facebook does not allow downloading of data from the groups.

Ensuring the privacy of social media users is recognised as a critical issue in social media research. Even if the data is publicly available, it does not mean that it can be used without considering privacy protection. On the contrary, Zimmer (2010) argues that 'privacy and anonymity do not disappear simply because subjects participate in online social networks; rather, they become even more important'.

In this case study, the observation of the group was agreed upon in a discussion with the group moderators. Furthermore, the researcher posted a message to the group describing how the observation was done. It was emphasised that no personal information (names or quotes) would be published without asking for permission. Observation was done afterwards by manually browsing the group news feed.

4.6. Map Analysis

The crowdsourcing data was manually analysed with a qualitative approach. The data collected, i.e., results from the map questionnaire and posts from social media, were brought together on one map. Combining all the different data sources was laborious, largely because the social media posts were not geotagged, and the locations of the posts were marked manually on the map if it was possible to identify the place. The locations of the images and posts were not recognisable in all posts, and, thus, it was impossible to show all the data on the map: altogether, only 68 out of 191 Instagram posts (i.e., 36%) were included in the map analysis. Despite this, it seems that social media do support and complement crowdsourcing with the questionnaire: for example, some demolished buildings (e.g., the old bus station in Figure 5, gas station and milk central) are mentioned in both forms of data, but pictures of those buildings are shared only in social media.

The social media also tell stories not available in the questionnaire data. For example, buildings and sites that are currently under construction are emphasised in social media even if the buildings have not been recognised as important. This naturally indicates the importance of the local environment and the ongoing change in the town.

The conjunctive map analysis revealed various findings (numbers indicate the locations on the map in Figure 4):

- Demolished buildings (1) are part of the cultural heritage of Nikkilä. These buildings are mentioned

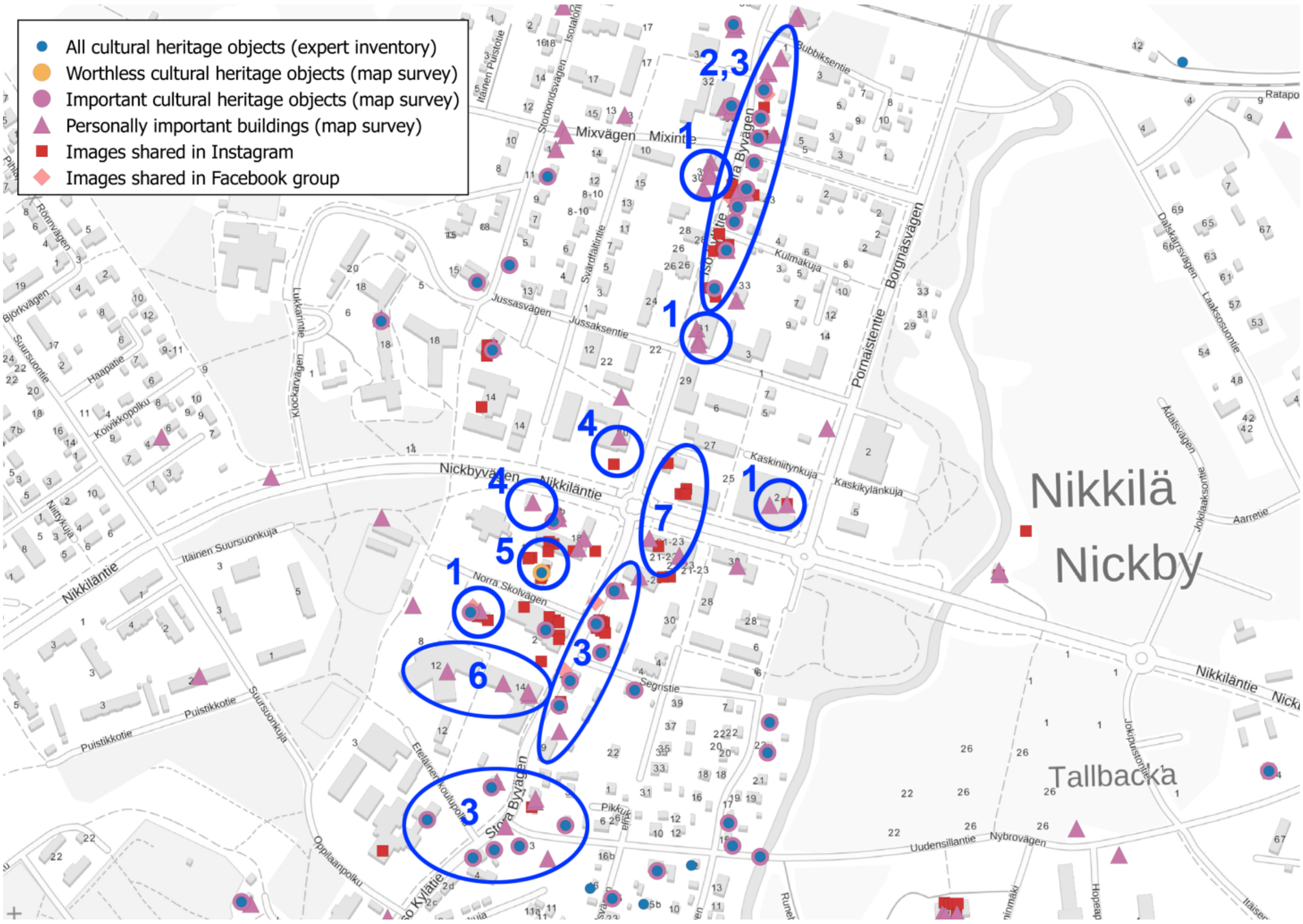


Figure 4. Thematic map of crowdsourced data from map questionnaire and Instagram. Copyright: Map background: National Land Survey of Finland, 10 November 2017.

in both the questionnaire and social media, but pictures of these are shared only in social media.
- Buildings that are evaluated as important in the PPGIS data also get attention in social media (such as the old wooden houses on Iso Kylätie) (2).
- In the centre of Nikkilä, most of the cultural heritage objects evaluated as valuable by experts in the cultural heritage inventory of Sipoo are also valued as personally important by the respondents (3).
- Some personally valued buildings are not considered important in the cultural heritage inventory (for example, an old barn and a restaurant from the '70s) (4).
- There is only one cultural heritage object, an old apartment block from the 1940s, that was evaluated as invaluable in PPGIS (5). However, this building was recognised many times in social media, because of an art work: a mural painted on the walls of the building.
- Meeting places such as schools, grocery stores and restaurants are often evaluated as personally important buildings in the PPGIS data (6).
- In social media, buildings and sites that are currently under construction are emphasised even if the buildings have not been recognised as important in PPGIS (7).

5. Results from the Case Study Nikkilä Memories

The crowdsourcing model used in the case study combined different data sources: map questionnaire, Instagram and Twitter posts and observation of a self-organized Facebook group. The data sources appeared to be powerful sources of three different types of local knowledge: PPGIS provided knowledge for comparison of expert and non-expert values and place-based memories; social media (especially Instagram) provided a tool for gathering and sharing visual representations of place-based experiences; the self-organized Facebook group appeared to be a tool for collective memory and knowledge building. In the following sections, the results are discussed in detail.

5.1. Functionality of the Crowdsourcing Model

In this article, the main research focus is on the crowdsourcing model and how functional the combination of PPGIS and social media crowdsourcing was in this case. The results indicate that this kind of multi-method crowdsourcing is beneficial but laborious to implement and analyse. It is evident that the methods complement each other; with PPGIS, it was possible to evaluate the importance of old buildings and collect textual place-based memories. Social media, especially Instagram, on the other hand, provided a large amount of visual representations of people's experiences from Nikkilä. However, it seems difficult to capture the authentic experience from the Instagram posts. In some cases, users describe their feelings about the places but often the interpretation remains open. Apart from that, only a fraction of the images were geotagged and the monitoring tool that was used did not collect location data. Thus, the Instagram data presented on the analysis map was geotagged manually by the researcher. To do so, it was necessary to be very familiar with the place.

The Instagram data consisted mostly of instant experiences rather than memories. Exceptions were a couple of historical images, one representing the old bus station (Figure 5) and a landscape photo from the beginning of

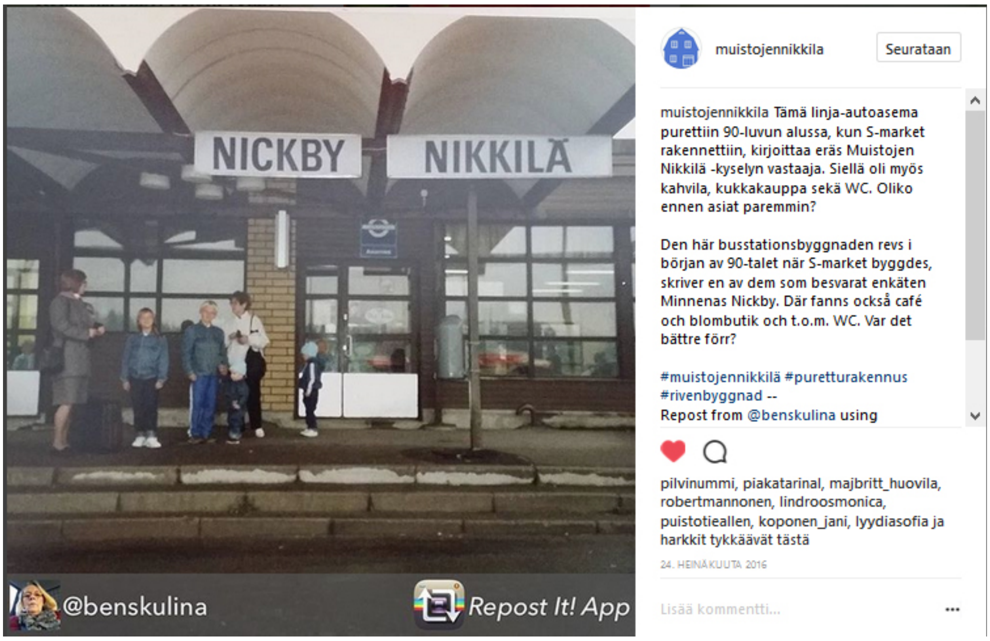


Figure 5. Nikkilä bus station, image shared on Instagram. Copyright: Benita Christina Lipponen. Used with permission.

the twentieth century. Naturally, it requires more motivation to post an old picture to Instagram since the tool is designed for posting images taken with mobile device.

5.2. Comparison between Local Knowledge and Expert Knowledge

The map analysis (Figure 4) provides an insight into residents' views in relation to expert knowledge by visualising the cultural heritage buildings identified by experts and buildings valued by residents in parallel. It shows that, in the centre of Nikkilä, almost all the cultural heritage buildings defined by the authorities are appreciated also by the respondents of the questionnaire. Only one building in the centre of Nikkilä, a former apartment building built in the 1940s (Figure 6), was evaluated negatively by the respondents due to its bad condition and 'ugly' appearance. The results indicate a reasonable level of consensus about the values of the built cultural heritage between the experts and local people. Yet, there are buildings that were reported as valuable personally for the respondents, but are not recognised as culturally valuable by the authorities. These buildings are, for example, schools, stores, homes and locations of buildings that have been demolished, such as the old bus station of Nikkilä, two gas stations and kiosks where kids used to buy candies.

Pictures of buildings that were considered valuable in the questionnaire were also shared on social media. Hence, it seems that social media complement and confirm the results from the questionnaire. On Instagram, people shared pictures of personally important buildings, but also new buildings that were under construction in the centre of Nikkilä. In this way, the change of the physical environment is documented on social media.

5.3. The Value of Place-Based Memories: Revealing the Intangible Cultural Heritage

Based on qualitative analysis of this heterogeneous data set, it was possible to find out how people value cultural heritage buildings in Nikkilä. Furthermore, based on people's memories and experiences presented in the PPGIS and social media, it was possible to understand the reasons why these buildings are valued. For example, many childhood memories from the mid-twentieth century were reported. This rich source of memories and experiences can be seen as a representation of intangible cultural heritage: the data reveals the local history of Nikkilä as experienced by local people and also reveals places that are currently provoking new memories.

The map questionnaire data, especially the memories mapped to the old wooden houses along the main street, Iso Kylätie, reveal the intangible aspects of cultural heritage. A lot of memories are related to shops and services that used to operate in the wooden houses along the main street of the village, Iso Kylätie. For example, these quotes describe the memories related to shops and services (free translations by the author):

> As a child, I went shopping in "Ässä" many times. I have danced until the early hours in "Kellari".
>
> Was it here where Lagerqvist was selling cheese? At some point there was an electrical shop on one side of the building also.
>
> There was Broström's car spare part shop in this building. Earlier he had a gas station. There was everything you needed.
>
> Rosenholm's was a legendary shop where you could buy everything you needed. The smell in the shop was extremely fine and service was always good and

Figure 6. An old apartment building called 'Aravatalo' in the centre of Nikkilä was evaluated as worthless. Copyright: author.

> friendly! My mother used to work there as an assistant and that felt extremely fabulous as a child.
> This is where we used to shop at the beginning of the 1950s when we were riding bikes. Rosenholm had everything, all the possible spare parts for bicycles. Books, bicycles, mopeds, LPs, fabrics, threads and nails etc. Here I bought my Billnäs spanner to be able to change tyres myself.

Some of these old buildings still exist and new local services, like a café, a decoration shop, a car service and a flea market, operate in the houses. Based on the results of the crowdsourcing, it is evident that this local 'shopping street' is an important part of Nikkilä's intangible cultural heritage, and developing it further as an active part of Nikkilä would support the local identity.

It is not surprising that a local shopping street with a variety of services and shops is a central part of people's memories. Zukin (2012) has pointed out in a case study in Amsterdam that local shopping streets shape the intangible cultural heritage and store collective memory. She emphasises that modern consumer culture and globalised businesses are a threat to these streets, and, therefore, official protection is needed. In Nikkilä, the intangible culture of the local businesses and old buildings on the main street is in danger of extinction.

5.4. Building a Collective Memory in Self-Organized Facebook Group

The local self-organized Facebook group 'Old buildings in Sipoo' appeared to be a good source of local history and a place for shaping a collective memory; people seemed to be eager to share information, memories and pieces of local history there. It is quite obvious that in a local group specialising in old buildings, people discuss the history of buildings and places, and, thus, they build a collective memory together.

Sharing memories in the Facebook group often started when a picture of an old building or a historical picture was posted to the group. This makes it challenging or even impossible to automatically monitor memories in social media. In this case, discussions in the Facebook group were observed manually by following the group as a member. In fact, on the record, there are no social media monitoring tools that can be used for monitoring closed Facebook groups.

One example of an image that evoked memories is the picture of old wooden houses on the main street of Nikkilä (Iso Kylätie) (Figure 7). The original post only consists of the year (presumably the year the image was taken) and the photograph. Inspired by the photo, people started to discuss what kind of shops there have been in the buildings and what they had bought (or wished to buy) there. For example, in this case the discussion starts with a memory of a book shop and continues with the items they were selling (freely translated by the author):

> There was Rosenholm's book and gift store in the blue house. (Member 1)
> Rosenholm had much more: toys, fireworks, fabrics, yarns, sewing material, bikes and repair and spare part service for bikes and mopeds. (Member 2)
> We had our eyes on the toys in that gable window. (Member 1)

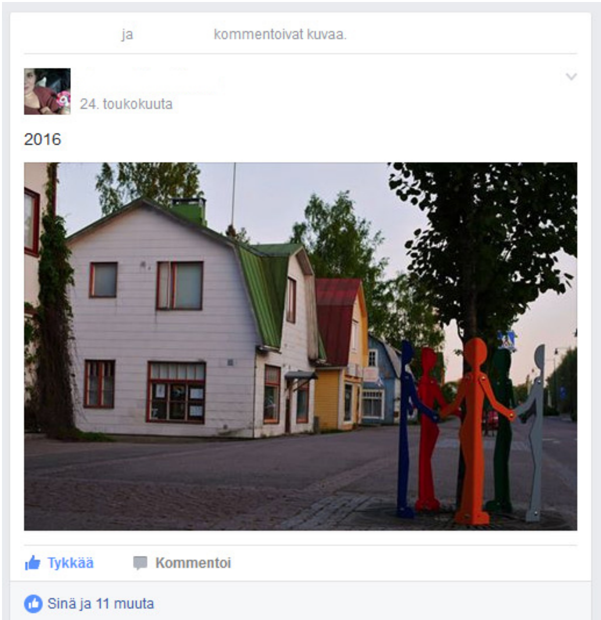


Figure 7. Screenshot from Facebook group 'Old buildings in Sipoo'. An image that evoked memories. Copyright of the photograph: Jenna Seppänen. Used with permission.

This kind of information describing the use of the building is useful for cultural history inventories; it describes the intangible part of cultural history that is often missing. Apart from that, these results suggest that people's memories should be a part of the knowledge base of urban planning. Memories and stories are easy to empathise with, and it would help planners to see the place through other people's eyes and to better understand the intangible cultural heritage.

As on Instagram, the buildings that were commented on Facebook are mostly the same as those that were evaluated as important in PPGIS. The difference on Facebook is that people also share their memories and discuss the local history. From a crowdsourcing perspective, it is disappointing that such data is not open to be used in urban planning.

5.5. Crowdsourcing Data—Benefits and Challenges for Planning Practices

The results show how crowdsourcing with different tools provides different kinds of data. All the tools used in the crowdsourcing have their advantages and flaws. The main findings are presented in Table 2.

Table 2. Comparison of crowdsourcing methods.

	PPGIS	Instagram	Facebook group
Data	186 respondents	191 posts shared with hashtag #muistojennikkilä	13 discussion threads related to built cultural heritage in Nikkilä
Users/respondents	Residents of Nikkilä, other parts of Sipoo or former residents of Nikkilä	126 followers, active users (who post images) are local residents from Nikkilä	Local residents interested in old buildings in Sipoo
Benefits	Geotagged data Comparison of experts' and residents' viewpoints	Visual representations complement PPGIS data	Enables building a collective memory
Challenges	Laborious analysis: GIS is not designed for qualitative content analysis	Availability of the data: collecting data needs a monitoring tool, the tool does not collect geodata, geo-locating the posts requires extensive familiarity with the place	Availability of the data: it is not allowed to download data from a closed Facebook group

As mentioned before, analysing the heterogeneous crowdsourcing data was laborious. Three main challenges were found during the analysis process: first, processing the qualitative GIS data from the questionnaire is not a straightforward process; GIS applications are not designed for qualitative analysis. Second, social media data was not geotagged. Since June 2016, Instagram does not provide an open API that would allow a researcher to crawl and download user-generated data. Therefore, to be able to perform a comprehensive analysis on the map, social media data was manually geotagged. Third, planners are still used to report-based practices. It became evident in the analysis phase that planners wanted the PPGIS data to be reported as a document where each building with the memories related to it is presented separately. The other option would have been to develop a map tool that combines all crowdsourcing data on one interactive map. This observation is in line with Eräranta and Staffans (2015), who argue that urban planning is still a strongly report-based activity. Therefore, integrating crowdsourcing into urban planning requires new data-oriented planning practices instead of working with static reports. In fact, this is crucial, as in practice it is impossible to publish user-generated social media data as a report document due to the copyrights.

5.6. The Reflection to the Planning of Nikkilä

In the development plan of Nikkilä (Municipality of Sipoo, 2016), the planners of Sipoo analysed this information and translated it to alignments that guide future detailed planning. In that plan, it was, for example, identified that old milieus (especially the old wooden houses in Nikkilä) are not only dear to the residents, but also cornerstones of the identity of Nikkilä. Thus, it was aligned in the development plan that architectural heritage should be taken as a starting point when developing new building (Municipality of Sipoo, 2016). This indicates that, in this case, buildings significant to participants should be maintained, and the new building should be fitted to the cultural heritage. However, it is mentioned in the development plan that the characteristics of traditional building can also be expressed by contemporary architecture. This means that the aim is to develop a townscape that contains different architectural layers that reflect the building period: old buildings with traditional style are preserved and new buildings with contemporary architecture complement the cultural heritage.

Currently, an outline plan and several detailed plans for the Nikkilä area are in the making. In the outline plan, the planners generated user profiles that represent different kinds of residents in Nikkilä. These profiles are then used for evaluating planning solutions. The planners also mentioned that the data can be used as background information for different future planning projects. The aim is to continue this study with a follow-up phase to find out how the data are used in these ongoing planning projects.

The analysis presented in this article, produced by the researcher, provides knowledge for the outline planning process of Nikkilä. It provides an interpretation of the sociocultural heritage of Nikkilä, a presentation that can and should be considered when evaluating the planning solutions. Especially considering how commercial services are developed and where are they located in the city structure is an important question. Also, how will traffic flows affect the old shopping street of Nikkilä: does it help business to flourish or not?

6. Discussion

Regardless of the encouraging results, this approach has several shortcomings in relation to representativeness of the participants, the useability of the PPGIS questionnaire, availability of social media data, difficulties in analysing heterogeneous data sets and limited usefulness of the results for urban planning.

The limitations of the PPGIS method relates to the representativeness of the participants. It is evident that the selection of tools using social media and internet as a primary marketing channel ended up with a result where respondents do not represent all age groups. Especially older respondents (over 66 years) are underrepresented, but also younger age groups (under 25 years) are missing based on the questionnaire results. Considering the number of residents in Nikkilä, the PPGIS method provided a relatively good number of respondents (in total, 186). However, in order to gather a comprehensive data set for evaluating the value of the old buildings, a larger respondent group would have been needed. Also, it became evident that there were useability problems with the questionnaire; some respondents claimed that it was too difficult to use the map and evaluating the old buildings was burdensome.

The PPGIS method was designed primarily for comparison of local and expert opinions. It is probable that presenting the cultural heritage objects valued by experts as a basis for the evaluation had an influence on the responses. In the questionnaire, respondents were asked to evaluate buildings considered valuable by the experts, but, after this stage, they were given a possibility to add personally important buildings on the map. Altogether, 106 personally important buildings were mapped, of which one proportion was old buildings already demolished, one proportion included some buildings that were presented in the expert evaluation, and one proportion were buildings considered important by respondents but not valued by the experts. If the goal of the case study would have been only to identify personally important buildings, it would have been better to leave out the expert evaluation data and let the respondents map important buildings from the scratch, which would probably have ended up as a different result.

One significant bias identified in the quality of the memories was that negative memories were missing. There were no memories, for example, about war, unemployment or poverty, even if, in the history of Nikkilä, there are relations to Finnish military history. The reason for this bias is not clear; maybe the questionnaire could have been designed differently to encourage looking back to negative things as well.

As mentioned in the results, the heterogeneous data set was laborious to analyse as the crowdsourcing model was new and no existing analysis tools were available. The map analysis required manual mapping of the social media data, and the qualitative analysis of PPGIS data was done manually as well.

For urban planning, this method provided a new type of knowledge, and, beforehand, it wasn't clear how to use the data. During this study, it became evident that the usefulness of the data is limited, and that planners prefer to use reports instead of raw data in planning. In the future, the aim is to study further how the data and results of this study will be used in the planning of Nikkilä.

With the use of social media, the problems relate to the copyright and use right issues. For example, it is not allowed to publish images shared on Instagram without owner's permission, and, thus, using the shared images in planning is not straightforward. Additionally, downloading data from closed Facebook groups is against Facebook's rules, and, therefore, the value of the collective memories generated on Facebook remains limited for urban planning. Furthermore, the images shared on Instagram were actually rather representing instant experiences than place-based memories. It seems to be difficult to facilitate social media users to answer specific questions. Nevertheless, it is evident that facilitating the content production on Instagram affected the users' activities; selected images were shared further (reposted) by the Nikkilä Memories profile, and, as users consider reposting of their content as a reward, that may have had an effect on the content people share.

7. Conclusions

This study contributes to research by developing and testing a crowdsourcing method that combines PPGIS and social media. In the case study, a rarely used type of local knowledge—place-based memories—is applied to the planning of a small, but growing and densifying town that has cultural heritage values. Experiences from the case study are especially relevant in cases where a culturally sensitive planning approach is chosen. Furthermore, results provide knowledge for further development of the crowdsourcing model and methods that aim to compare local knowledge and expert knowledge.

The results indicate that crowdsourcing is a functional tool for gathering place-based memories and revealing local people's values related to old buildings and intangible cultural heritage, and comparing local knowledge with expert knowledge. Furthermore, it is apparent that social media data can complement other forms of data such as PPGIS data. Similar findings have been identified in other studies as well (e.g., Heikinheimo et al., 2017). In this study, social media data complemented PPGIS data with visual representations of memories and experiences, and provided local people a place for shaping collective memories.

As Pánek and Benediktsson (2017) argue, systems that support participatory planning can help to bridge the gap between planners and the public. In this case, the gap between expert knowledge and local knowledge was bridged with a crowdsourcing model in three different ways. First, local people's values in relation to built cultural heritage were studied by comparing experts' and citizens' evaluations of old buildings in Nikkilä. Second, the intangible dimensions of Nikkilä's cultural heritage were identified to help planners understand the identity of Nikkilä. Third, by developing resident profiles to be used for evaluating planning solutions, the planners used the data for understanding local viewpoints.

This study revealed several challenges for using a multi-method crowdsourcing model in collaborative planning. As mentioned earlier, participatory planning is an interactive process (Healey, 1997). It is arguable that the process of Nikkilä Memories has some interactive elements, and especially the use of social media added interaction to the process. However, using the PPGIS tool is mostly non-interactive in the sense that it doesn't create a dialogue between people and planners. Thus, from the participatory planning perspective, it is relevant to continue the process with interactive planning methods.

Collaborative knowledge building that aims at shared understanding is a central element of participatory urban planning (Healey, 1999). Planning theories do not provide tools for knowledge building in practice, and, therefore, empirical studies that solve the challenges of collecting, analysing and managing user-generated data are needed. In this study, a closed Facebook group appeared to be a functional tool for knowledge building, as people were sharing and discussing their memories there and created collective memories. However, there is still a need for new tools to share the crowdsourcing results with a wider audience.

As Eräranta and Staffans (2015) argue, it seems that urban planning still is a largely report-based activity. It became evident during the crowdsourcing project that planners prefer static reports instead of using data with interactive applications. An interactive web map was produced but still the planners required the results as a report. These findings underline the need to develop data-oriented practices for urban planning. Moreover, there is a need to develop a useable and efficient analysis method for heterogeneous user-generated data. Especially the use of social media data adds challenges to data analysis. From the crowdsourcing perspective, it is also relevant to consider methods that allow local people not only to participate as knowledge producers, but also to engage people in idea generation, as argued by Brabham (2009).

Acknowledgements

I want to thank all the local people in Sipoo who took part in Nikkilä Memories project by providing their memories and values in the crowdsourcing. Without active citizens this project would not have been successful. The municipality of Sipoo has provided funding for the map survey tool used in the case study. Many thanks to general planning architect Eveliina Harsia and urban planner Aino Kuusimäki for making this project possible and cooperation in the project. Also, many thanks to Maarit Kahila and Anna Broberg from Mapita Ltd who assisted the project team in developing the map questionnaire.

Conflict of Interests

The researcher has worked as a town planning architect in the Municipality of Sipoo.

References

Aiello, L. M., Schifanella, R., Quercia, D., & Aletta, F. (2016). Chatty maps: Constructing sound maps of urban areas from social media data. *Royal Society Open Science*, *3*(3), 1–19.

Bianchini, F., & Ghilardi, L. (2007). Thinking culturally about place. *Place Branding and Public Diplomacy*, *3*(4), 280–286.

Brabham, D. C. (2009). Crowdsourcing the public participation process for planning projects. *Planning Theory*, *8*(3), 242–262. https://doi.org/10.1177/1473095209104824

Brown, G. (2015). Engaging the wisdom of crowds and public judgement for land use planning using public participation geographic information systems. *Australian Planner*, *52*(3), 199–209. https://doi.org/10.1080/07293682.2015.1034147

Campagna, M. (2014). The geographic turn in social media: Opportunities for spatial planning and geodesign. In B. Murgante, S. Misra, A. M. A. C. Rocha, C. M. Torre, J. G. Rocha, M. I. Falcão, . . . O. Gervasi (Eds.), *Computational science and its applications–ICCSA 2014* (pp. 598–610). Berlin: Springer.

Campagna, M., Floris, R., Massa, P., Girsheva, A., & Ivanov, K. (2015). The role of social media geographic information (SMGI) in spatial planning. In S. Geertman, J. Ferreira Jr., R. Goodspeed & J. Stillwell (Eds.), *Planning support systems and smart cities: Lecture notes in geoinformation and cartography* (pp. 41–60). Basel: Springer International Publishing Switzerland.

Cranshaw, J., Schwartz, R., Hong, J., & Sadeh, N. (2012). The Livehoods project: Utilizing social media to understand the dynamics of a city. In *Proceedings of the International AAAI Conference on Web and Social Media* (pp. 58–65). Dublin: AAAI Press.

Dunkel, A. (2015). Visualizing the perceived environment using crowdsourced photo geodata. *Landscape Architecture and Environmental Planning. Special Issue: Critical Approaches to Landscape Visualization*, *142*, 173–186.

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, https://doi.org/10.1080/00045608.2011.595657

Eräranta, S., & Staffans, A. (2015). *From situation awareness to smart city planning and decision making*. Paper presented at the 14th International Conference on Computers in Urban Planning and Urban Management, July 7–10, 2015, Cambridge, MA.

Facebook. (2017). What are the privacy settings for groups? *Facebook Help Center*. Retrieved from www.facebook.com

Finlex. (1999). *Maankäyttö- ja rakennuslaki* [Land Use and Building Act]. Retrieved from www.finlex.fi/fi/laki/ajantasa/1999/19990132

Frias-Martinez, V., & Frias-Martinez, E. (2014). Spectral clustering for sensing urban land use using Twitter activity. *Engineering Applications of Artificial Intelligence*, *35*, 237–245. https://doi.org/10.1016/j.engappai.2014.06.019

Healey, P. (1997). *Collaborative planning. Shaping places in fragmented societies*. Hampshire and London: Macmillan Press.

Healey, P. (1999). Institutionalist analysis, communicative planning, and shaping places. *Journal of Planning Education and Research*, *19*, 111–121.

Heikinheimo, V., Di Minin, E., Tenkanen, H., Hausmann, A., Erkkonen, J., & Toivonen, T. (2017). User-generated geographic information for visitor monitoring in a national park: A comparison of social media data and visitor survey. *ISPRS International Journal of Geo-Information*, *6*(3), 85; https://doi.org/10.3390/ijgi6030085

Howe, J. (2006, June). The rise of crowdsourcing. *Wired Magazine*. Retrieved from www.wired.com

Hu, Y., Gao, S., Janowicz, K., Yu, B., Li, W., & Prasad, S. (2015). Extracting and understanding urban areas of interest using geotagged photos. *Computers, Environment and Urban Systems*, *54*, 240–254.

Kahila, M., & Kyttä, M. (2008). SoftGIS as a bridge builder in collaborative urban planning. In S. Geertman & J. Stillwell (Eds.), *Planning support systems: Best practices and new methods*. Springer.

Kahila-Tani, M., Broberg, A., Kyttä, M., & Tyger, T. (2016). Let the citizens map—Public participation GIS as a planning support system in the Helsinki Master Plan process. *Planning Practice & Research*, *31*(2), 195–214.

Khan, G. (2015) Social media-based government explained: Utilization model, implementation scenarios, and relationships. In I. Boughzala, M. Janssen, & S. Assar (Eds.), *Case studies in e-government 2.0*. Springer, Cham.

López-Aparicio, S., Vogt, M., Schneider, P., Kahila-Tani, M., & Broberg, A. (2017). Public participation GIS for improving wood burning emissions from residential heating and urban environmental management. *Journal of Environmental Management*, *191*, 179–188. https://doi.org/10.1016/j.jenvman.2017.01.018

López-Ornelas, E., & Zaragoza, N. M. (2015). Social media participation: A narrative way to help urban planners. In G. Meiselwitz (Ed.), *Social computing and social media, LNCS* (Vol. 9182, pp. 48–54). Basel: Springer International Publishing Switzerland.

Mueller, J., Lu, H., Chirkin, A., Klein, B., & Schmitt, G. (2018). Citizen design science: A strategy for crowd-creative urban design. *Cities*, *72*(A), 181–188. https://doi.org/10.1016/j.cities.2017.08.018

Municipality of Sipoo. (2006). *Sipoon kunnan kulttuuriympäristö- ja rakennusperintöselvitys* [Inventory of cultural heritage in the Municipality of Sipoo]. Arkkitehtitoimisto Lehto Peltonen Valkama Oy, Ympäristötoimisto Oy.

Municipality of Sipoo. (2016). *Nikkilän kehityskuva* [The development plan for Nikkilä]. *Nikkilä Nickby*. Retrieved from www.nikkilanickby.fi

Nakatsu, R. T., Grossman, E. B., & Charamblos, L. I. (2014). A taxonomy of crowdsourcing based on task complexity. *Journal of Information Science*, *40*(6), 823–834. https://doi.org/10.1177/0165551514550140

Nielek, R., Ciastek, M., & Kopeć, W. (2017). *Emotions make cities live. Towards mapping emotions of older adults on urban space*. Conference paper in WI '17, August 23–26, 2017, Leipzig, Germany.

Nummi, P. (2017). Social media data analysis in urban e-planning. *International Journal of E-Planning Research*, *6*(4), 18–31.

Nummi, P., & Tzoulas, T. (2015). Engaging citizens in cultural planning with a web map survey. *Culture and Local Governance*, *5*(1/2), 161–173.

Official Statistics of Finland. (2017). Use of information and communications technology by individuals. *Statistics Finland*. Retrieved from www.stat.fi

Paldino, S., Bojic, I., Sobolevsky, S., Ratti, C., & González, M. C. (2015). Urban magnetism through the lens of geo-tagged photography. *EPJ Data Science*, *4*(5), 1–17.

Pánek, J., & Benediktsson, K. (2017). Emotional mapping and its participatory potential: Opinions about cycling conditions in Reykjavík, Iceland. *Cities*, *61*, 65–73.

Quercia, D., Aiello, L. M., & Schifanella, R. (2016). The emotional and chromatic layers of urban smells. In *Proceedings of the 10th International AAAI Conference on Web and Social Media (ICWSM)* (pp. 309–318). California, Palo Alto: AAAI Press.

Resch, B., Summa, A., Zeile, P., & Strube, M. (2016) Citizen-centric urban planning through extracting emotion information from Twitter in an interdisciplinary space-time-linguistics algorithm. *Urban Planning*, *1*(2), 114–127. https://doi.org/10.17645/up.v1i2.617

Seltzer, E., & Mahmoudi, D. (2012). Citizen participation, open innovation, and crowdsourcing: Challenges and opportunities for planning. *Journal of Planning Literature*, *28*(1), 3–18. https://doi.org/10.1177/0885412212469112

Sykora, M., Robertson, C., Shankardass, K., Feick, R., Shaughnessy, K., Coates, B., . . . Jackson, T. W. (2015). Stresscapes: Validating linkages between place and stress expression on social media. In I. Katakis, F. Schnitzler, T. Liebig, D. Gunopulos, K. Morik, G. Andrienko, & S. Mannor (Eds.), *Proceedings of the 2nd International Workshop on Mining Urban Data* (pp. 80–84).

Tasse, D., & Hong, J. I. (2014). Using social media data to understand cities. In P. Thakuriah, N. Tilahun, & M. Zellner (Eds.), *Proceedings of NSF Workshop on Big Data and Urban Informatics* (pp. 64–79). Chicago: University of Illinois Press.

Van Herzele, A. (2004). Local knowledge in action. Valuing nonprofessional reasoning in the planning process. *Journal of Planning Education and Research*, *24*, 197–212.

Yang, P., & Ng, T. L. (2017). Gauging through the crowd: A crowd-sourcing approach to urban rainfall measurement and storm water modeling implications. *Water Resources Research*, *53*(11), 9462–9478. https://doi.org/10.1002/2017WR020682

Zhou, X., & Zhang, L. (2016). Crowdsourcing functions of the living city from Twitter and Foursquare data. *Cartography and Geographic Information Science*, *43*(5), 393–404. https://doi.org/10.1080/15230406.2015.1128852

Zimmer, M. (2010). "But the data is already public": On the ethics of research in Facebook. *Ethics and Information Technology*, *12*(4), 313–325.

Zukin, S. (2012). The social production of urban cultural heritage: Identity and ecosystem on an Amsterdam shopping street. *City, Culture and Society*, *3*, 281–291.

About the Author

Pilvi Nummi (architect MSc) is a doctoral candidate in the Department of Built Environment, School of Engineering, at Aalto University (Finland). Her research interest is in e-Planning and e-Participation in urban and land-use planning, especially in the use of social media in participatory urban planning. Since year 2008, she has worked as a town planning architect in Finland.

Urban Planning (ISSN: 2183-7635)

Urban Planning is an international peer-reviewed open access journal of urban studies aimed at advancing understandings and ideas of humankind's habitats.

www.cogitatiopress.com/urbanplanning

COGITATIO