Research Paper

Social media and the city: Rethinking urban socio-spatial inequality using user-generated geographic information

Taylor Shelton\textsuperscript{a,}\textsuperscript{*}, Ate Poorthuis\textsuperscript{b}, Matthew Zook\textsuperscript{b}

\textsuperscript{a} Clark University, Graduate School of Geography, 950 Main Street, Worcester, MA 01610, United States
\textsuperscript{b} University of Kentucky, Department of Geography, Lexington, KY, United States

HIGHLIGHTS

• Analyzes two years of geotagged tweets from Louisville, Kentucky.
• Explores popular spatial imaginaries of the '9th Street Divide'.
• Argues for greater linkages between socio-spatial theory and big data research.
• Develops a novel conceptual and methodological frame for using social media data.

ABSTRACT

Big data is increasingly seen as a way of providing a more 'scientific' approach to the understanding and management of cities. But most geographic analyses of geotagged social media data have failed to mobilize a sufficiently complex understanding of socio-spatial relations. By combining the conceptual approach of relational socio-spatial theory with the methods of critical GIScience, this paper explores the spatial imaginaries and processes of segregation and mobility at play in the notion of the '9th Street Divide' in Louisville, Kentucky. Through a more context-sensitive analysis of this data, this paper argues against this popular spatial imaginary and the notion that the Louisville’s West End is somehow separate and apart from the rest of the city. By analyzing the everyday activity spaces of different groups of Louisvillians through geotagged Twitter data, we instead argue for an understanding of these neighborhoods as fluid, porous and actively produced, rather than as rigid, static or fixed. Ultimately, this paper is meant to provide a conceptual and methodological framework for the analysis of social media data that is more attentive to the multiplicity of socio-spatial relations embodied in such data.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Attempts to make the ideas and practices of urban planning more ‘scientific’ through the application of new technologies have been persistent over the course of the last century (Fairfield, 1994; Ford, 1913; LeGates, Tate, & Kingston, 2009; Light, 2003). But as new sources of digital data – whether collected from mobile phones, social media feeds, sensors embedded in the built environment or any number of other sources – are increasingly able to be combined and cross-referenced to produce ‘big data’, there has been a revival of interest in mobilizing this data toward the end of a supposedly more holistic ‘science of cities’ (Bettencourt & West, 2010; Batty, 2012). The breadth of available data sources has expanded rapidly, allowing researchers to end their dependence upon official statistics on demographics, economic activity, traffic, and any number of other urban indicators.

But as these new data sources and new ways of approaching social science research have become more prominent, they have also faced increasing amounts of criticism. This is due in no small part to the hubris of big data advocates, as exemplified in Anderson’s (2008) now-infamous declaration of the ‘end of theory’. And while claims to greater objectivity, neutrality, and accuracy are rampant among proponents of big data, boyd and Crawford (2012) astutely argue that these data are always the result of conscious, subjective decisions on the part of researchers, and are the result of inherently social processes. Indeed, it is important to keep in mind that in spite of the celebratory discourses around big data, many of these ideas and techniques have been around for considerable amounts of time (Barnes, 2013; Graham & Shelton, 2013; Wyly (2014); however, positions big data as the driver of a ‘new
quantitative revolution’ in geography, a largely reductionist effort enabled by processes of neoliberalization which threaten the kind of situated research which geographers have become experts at producing. In describing what he sees as “the speedy pseudopositivism of tweet-space analysis”, Wyly argues that “big data give us a quickly expanding, shallow view of the vast horizontal landscape of the desert of the present real, with each new technological advance accomplishing new kinds of devalorization of past generations of human knowledge” (Wyly, 2014:28).

While we are sympathetic to such critiques of big data, we also recognize that these traits are not inherent in the data themselves, nor in the analysis of such data. The use of big datasets is not necessarily reductionist or ahistorical; these are, in fact, to echo Wyly’s (2009) earlier analysis of quantification in geography, contingent circumstances. Indeed, we believe that big data can be quite easily fit into more critical-quantitative approaches to urban geography and planning (cf. Barnes, 2009; Schwanen & Kwan, 2009; Sheppard, 2001; Wyly, 2011). Though issues around the over-value-ization of this kind of data remain, including how they might displace other forms of official statistical knowledge, we believe there is also significant potential. For example, the finer spatial and temporal scale of these kinds of datasets provides a way to ask different kinds of questions than is possible with, for instance, census data, which is often several years old by the time of its release, and is generally associated only with one’s place of residence, and then aggregated to more-or-less arbitrary spatial units. As such, this paper highlights the potential in mobilizing big data sources for understanding urban socio-spatial processes, so long as such research is also explicit in its engagement with the appropriate conceptual and methodological frameworks, and built on a critical and contextualized understanding of the underlying data. When coupled with exactly the kind of historical and geographical context that Wyly sees missing from many big data analyses, we argue that these approaches can provide useful insight for urban planning and geographical research.

To this end, we use a dataset of geotagged tweets from Louisville, Kentucky to explore longstanding problems of socio-spatial inequality in the city. Louisville represents an interesting case study for a number of reasons: first, Louisville is something of an ‘ordinary’ city, especially when it comes to its reflection in these kinds of big data sources. Louisville is fairly average in the density of its social media footprint, meaning that the methods demonstrated in this paper are likely to be applicable to other localities, whereas a study of a New York City or another global outlier would beg the question of relevance for studying metropolitan areas more broadly. Second, Louisville is an increasingly prominent player in the landscape of data-driven urban governance, with Mayor Greg Fischer receiving national and international recognition for various policy initiatives aimed at making data, including data from social media platforms, a key driver in municipal policy development and implementation (Carroll, 2013; Fischer, 2012; Goldsmith, 2013; Louisville Metro Government, 2012; Reno-Weber & Niblock, 2013; Shelton, Zook, & Wiig, 2015). Third and finally, Louisville is a city with intense social inequalities and a keen appreciation how they are manifest spatially. This is seen most clearly in the notion of the “9th Street Divide”, which signifies the material inequalities and imaginative distance that separates the city’s predominantly poor and African–American neighborhoods in the West End from more affluent and predominantly white areas throughout the rest of the city (Crtucher, 2013).

As such, this study provides an opportunity to show how big data can be mobilized to produce alternative understandings of cities and urban processes. It is, however, important to acknowledge that our choice of case study is not accidental, and that the insights gleaned from our analysis rely on our own local knowledge and understandings of the city’s social dynamics, taken from our experiences living in and conducting a variety of research projects in Louisville. Our choice to highlight this is, however, much more than simply an acknowledgement of our own situatedness and biases; it is also an explicit attempt to counter the notion that meaningful insights about cities can be gleaned simply by ‘crunching the numbers’. Understanding urban socio-spatial processes requires more than massive amounts of data and clever software algorithms; it also necessitates grounded understandings of local history and culture, and the broader political-economic forces at play. Thus, our goal in this paper is to highlight the usefulness of combining the conceptual approaches of critical socio-spatial theory with new methodological approaches being utilized to understand big social media data.

2. Information technologies and the contemporary urban condition

Though the use of new sources of data and other new technologies are at the center of many contemporary urban policy initiatives, information technologies have long played a prominent role in the way that urban spaces are conceived, planned and enacted. This is especially true of mapping and geographic information technologies, whether in the form of hand-drawn maps or Google Maps mashups used to display data interactively on the web (Schein, 1993; Söderström, 1996). And while these technologies have evolved from early computer models and planning support systems toward more participatory and web-based approaches to GIS, the nascent ‘smart cities’ movement has begun to shift these technologies from desktop computers toward being embedded in the fabric of the city itself, allowing for a continuous collection and analysis of heterogeneous data streams meant to make urban systems operate more rationally and efficiently (Greenfield, 2013; Kitchin, 2014).

2.1. Urban analysis in the era of Web 2.0 and big data

One of the most powerful ways that information technology is shaping urban life in the 21st century is through the production of digital content – text, photos, videos, etc. – tied to particular locations on the earth’s surface. While the act of creating a geotagged tweet, posting a photo to Instagram, reviewing a restaurant on Yelp or ‘checking in’ to your favorite park on Foursquare may seem relatively mundane, these platforms and data sources are allowing for new ways of interacting with, and studying, cities (Arribas-Bel, 2014). As both Goodchild (2007, 2009) and Graham (2010) have argued, these platforms of data production offer unprecedented possibilities for codifying local knowledge about otherwise neglected places and making it widely accessible, even opening up the possibility for non-positivist epistemologies of mapping (Elwood & Leszczynski, 2013; Warf & Sui, 2010). These platforms not only allow for such local knowledge to be transferred to or accessed from distant places, but they also allow citizens in close proximity to one another to interact in a place-specific way through digital networks (Hardy, 2007).

While this data can be incredibly important for helping tourists navigate through unknown places using their smartphones and a combination of location-based applications, the significance of this data for the purposes of this paper is our capability of collecting, aggregating, mapping and analyzing this data to understand how these digital data shadows are intimately intermingled with offline, material geographies of everyday life. Geotagged social media data has been used to research topics ranging from linguistic and religious differences (Graham & Zook, 2013; Shelton, Zook, & Graham, 2012; Wall & Kirdnark, 2012; Watkins, 2012), to differences in the
places frequented by locals and tourists in different cities (Fischer, 2010; Poorthuis, 2010). Others have used this kind of data to rethink how we conceptualize and define neighborhoods or other spaces of social affiniti. For instance, Cranshaw, Schwartz, Hong, and Sadeh (2012) ‘Livehoods’ project used Foursquare check-ins to re-draw neighborhood boundaries based on similarities in user mobility, while Stefanidis et al. (2013) expand to the broader scale of the nation to understand the ‘polycentric communities’ formed by attention to and engagement with current events in far off places. Others have attempted to explain spatio-temporal variations in this content, with attention to processes such as the weekly movement of students leaving a college town in search of other entertainment options (Li & Shan, 2013; see also Li, Goodchild, & Xu, 2013). Kelley (2013) uses a mixed method analysis to understand not only the influence of social inequality on shaping the data shadows within particular urban neighborhoods, but also the different perceptions and experiences of place embodied in such content.

Despite the range of issues this data can be used to address, it is important to keep in mind that offline, material social processes, such as persistent social inequalities, continue to shape the data as we interact with it, never including everyone equally or in a representative fashion (Crutcher & Zook, 2009; Graham, Hogan, Straumann, & Medhat, 2014; Graham & Zook, 2011). Even those who create such content might be marginalized through the voices of other contributors, or through the automatic sorting of software algorithms that judge a particular comment to be of lower value (Zook & Graham, 2007). As such, it is important to recognize the limits of such technologies, as they approach only an incredibly shallow vision of democratization, if any (Haklay, 2013). Nonetheless, these massive streams of real-time social data are being incorporated into automated systems collecting information on energy use, traffic congestion and any number of other urban processes, which are then used to make decisions about both the day-to-day operations and long-term planning of the city. From smartphone apps collecting data on potholes in order to prioritize areas of need in infrastructure maintenance in Boston to the monitoring of African–American teenagers’ social media activity in Louisville, these kinds of applications are increasingly popular among local governments (Crawford, 2013; Leonard, 2014). That is, this data is not only interesting for its ability to shed light on relatively mundane geographic processes; it is also being used to directly shape the way we live in cities today.

2.2. Cities and social media beyond the geotag: Re-engaging socio-spatial theory

A key shortcoming to both scholarly and applied uses of this kind of data is, we believe, the failure to capture the broader range of socio-spatial processes that are embedded in the data. While the practice of geotagging only allows for these pieces of content to be tied to a single pair of latitude and longitude coordinates, this single point does little to reflect the variegated and polymorphous geographies of everyday life that this content represents (Crampton et al., 2013). Though many of the aforementioned studies have attempted to mobilize more complex understandings of space and socio-spatial processes, we believe a more direct engagement with longstanding theoretical approaches to be a fruitful way to push forward research into the geographies of social media data.

As such, this paper specifically seeks to integrate the concepts and disposition of what might broadly been termed ‘relational socio-spatial theory’ into our analysis. Taking off from Massey’s (1991) original conceptions of a ‘global sense of place’, this work attempts to conceive of space as networked, fragmented and processual, rather than as a kind of fixed container with defined boundaries and characteristics. From reconceptualizations of globalization (Amin, 2002) to a new focus on mobility as a fundamental, defining characteristic of contemporary life (Sheller & Urry, 2006), a key tenet of this approach has been an inversion of Tobler’s so-called ‘first law of geography’—that all things are related, but near things are more related than far things. Instead, relational approaches suggest that “we cannot assume that local happenings or geographies are ontologically separable from those ‘out there’” (Amin, 2002:386). By focusing on the social relations that recurrently produce space and are in turn influenced by it, rather than simply privileging proximity in absolute, Cartesian space, Amin argues that we can begin to see “a subtle folding together of the distant and the proximate” (Amin, 2007:103). As social processes are more and more spatially extensive, owing at least in part to the increasing prevalence of information and communication technologies, our spatial categories similarly need to evolve so as not to assume universal connections between social activities or processes and the locations on the earth’s surface at which they occur. And while much of this work has been produced with specific reference to cities and the urban, it has also been used with respect to the broader spatial scale of the region (Allen & Cochrane, 2007; Amin, 2004), as well as to the sub-urban scale, as evidenced in Massey’s original focus on her own neighborhood of Kilburn in London. Though much of this ‘relational turn’ sees itself as counterposed to the conventions of Marxist political economy approaches (cf. ongoing debates in Amin & Thrift, 2002; Brenner, Madden, & Wachsmuth, 2011; Macleod & Jones, 2007; McFarlane, 2011), an attempt to return to the work of Henri Lefebvre and his understanding of ‘planetary urbanization’ has provoked a similar tendency to view the urban ‘without an outside’, as unevenly stretched across the space of globe through networks and flows that have come to define global capitalism (Brenner, 2013; Brenner & Schmid, 2014). Despite this lingering conflict between ostensibly opposed epistemological and ontological positions, work by Jessop, Brenner, and Jones (2008) and McCann and Ward (2010) has attempted to highlight the compatibility of these relational and territorial approaches, and the potential for combining these insights in geographical analysis. Indeed, with respect to the analysis of social media data, Shelton, Poorthuis, Graham, and Zook (2014) have previously mobilized Jessop et al.’s Territory-Place-Scale-Networks framework as a means by which to account for the multidimensionality and polymorphous nature of socio-spatial relations. Ultimately, this kind of relational approach is useful in that it does not attempt to impose arbitrary limits on one’s analysis. By broadening the scope of one’s geographical imaginary to processes that span localities and to places quite distant in absolute terms, one can gain a greater appreciation for and understanding of the social and spatial context in which contemporary social activities are situated.

3. Methodology and data collection

In order to operationalize the orientation of relational socio-spatial theory with respect to social media data, we draw on methods from qualitative and critical GIScience, taking advantage of the power of both qualitative and quantitative analysis of this data, and the utility of mapping and geovisualization for communicating such analyses. In addition to these connections, this work also draws on work which seeks to understand urban inequalities not simply through official statistics related to race/ethnicity, income, education and other indicators geolocated to one’s official residence and aggregated to more-or-less arbitrary spatial units. Ahas, Silt, Jarv, Saluveer, and Tiru (2010) and Silt and Ahas (2014) have previously demonstrated the alternative possibilities for understandings of inequality based not on these more-or-less static statistics, but on measuring people’s movements through cities and identifying the places that they actually inhabit on
an everyday basis, and then understanding the deeper meanings behind such patterns (see also Kwan, 2013; Wong & Shaw, 2011).

In order to combine the attention to individuals as active, conscious producers of their own everyday lives, while also understanding the socio-spatial context that structures these activities and interactions (Hägerstrand, 1970; Pred, 1984), we developed a method that allows us to move between individual-level analysis of Twitter users and a neighborhood or area-based analysis that situates these users in the spatial context of sub-urban areas within the city of Louisville. To begin, we collected all geotagged tweets with exact latitude and longitude coordinates from within Louisville, Kentucky from late June 2012 to early July 2014, yielding a total of 5.7 million tweets (see Fig. 2a, which represents a 1% random sample of these tweets for the year 2013). While this kind of map is a useful starting point, such visualizations of ‘dots on a map’ do little to overcome the understanding of social media data being defined simply by its latitude and longitude coordinates. This is a trend recently pushed to its logical conclusion by what cartographer Kenneth Field (2014) has called ‘animated ectoplasm maps’ of geotagged tweets generated with easy-to-use online mapping tools, but which allow for very little substantive understanding of geographical processes. Rather than simply describing the presence or absence of data points, our goal is to explore how each individual data point might be associated with other data points through spatial or social proximity, as well as situated within both the immediate and broader spatial context in which it was created.

In order to do this, we refined our dataset by drawing boundaries around the West End and the East End, using the same classification as a popular local magazine article comparing the two areas (Crutcher, 2013; see Fig. 1), and selecting only those tweets within the two areas, yielding approximately 450,000 tweets in the West End and 1.1 million tweets in the East End. But simply noting that there is a relative dearth of tweeting in the West End only serves to affirm more general notions of a persistent digital divide between such predominantly poor and African–American neighborhoods and those predominantly affluent and white areas of the East End. So, in order to understand the everyday geographies of individuals living in, or at least spending considerable amounts of time in, these neighborhoods, we devised a step-by-step approach in which users are classified as ‘belonging’ to one of the two areas.

To meet this definition, users were required to satisfy two criteria. First, we identified those users expected to have created at least 40 geotagged tweets from within one of the two areas, based on a 10% random sample of tweets from within each neighborhood. This threshold is chosen to select relatively active Twitter users and to make sure ‘belonging’ is not based on only a few tweets that happen to be sent from within the neighborhood. Second, for users that satisfied the first criteria, we subsequently collected all of their tweets from within the city – both within and outside of their ‘home’ neighborhood – and further selected only those users with at least 50% of their geotagged tweeting activity originating within one of the two neighborhoods.

These criteria are meant to ensure the inclusion of only those users who predominantly and persistently tweet from one of these two areas, while excluding users with short-lived or minimally used accounts, as well as any one-time or infrequent visitors to one of the areas in question. The 50% threshold also ensures that no users are classified as being both West End and East End users. The end result is a dataset consisting of 703 users from the East End generating 274,338 geotagged tweets and 662 users from the
West End with a total of 398,432 geotagged tweets (see Fig. 2b and c). This dataset forms the foundation from which we are able to analyze aggregate patterns of socio-spatial mobility and segregation through a comparison of daily activity spaces of these two different groups.

In order to situate each data point within its broader spatial context and allow for comparisons between the selected neighborhoods, we aggregated each individual data point to a hexagonal grid composed of units roughly 1 km in diameter (Scott, 1985; Carr, Olsen, & White, 1992). A key part of this aggregation process was adjusting the absolute count of tweets in each hexagonal area to control for the influence of ‘power users’, so that the resulting visualization was more reflective of the entire user base (cf. Poorthuis & Zook, 2014). This decision reflects the social practices surrounding Twitter, in which a small group of users produce a disproportionately large amount of tweets; an issue that is compounded when examining specific subsets of tweets restricted by time or space. For instance, within the larger dataset, one user created 65 tweets from the area around 2nd and Market Streets in Louisville in one six hour period but never again tweeted from this area. Unadjusted, this activity would give equal weighting to each of these 65 tweets as to the tweets of individuals who travel regularly to this place, or individuals who only visit once but produce a much smaller amount of content. While this may be appropriate for some kinds of studies – particularly those focused on individual behavior – our focus on the broader social and neighborhood context led to our approach to only include a maximum of five randomly selected tweets per user in any given hexagon. In other words, an adjusted tweet count of 50 for a given hexagon signifies that there were at least ten different users tweeting from that area. The sample size of 5 was chosen to provide some indication of the strength of a user’s relation to that specific place while not diverging too far from the median value of one tweet per user per hexagon. This resulted in a total of 50,948 adjusted tweets from East End users, and 50,451 adjusted tweets from West End users. These adjusted counts, still aggregated to the hexagonal bins, serve as the basis for the calculation of the odds ratio measure seen in Fig. 3.

4. Visualizing urban socio-spatial inequalities using social media data

Louisville, Kentucky is, like nearly all American cities, still marked by the legacies of racial segregation. From discriminatory housing policies of the mid-20th century that forced black residents into the city’s West End neighborhoods, to ongoing struggles over the busing of school children, segregation and attempts to fight it have been at the center of Louisville’s historic development (Blum, 2006; Cummings & Price, 1997; Poe, 2013). Ultimately, these structural forces produced a city increasingly divided

---

It is also worth noting that these struggles in Louisville have played a key role in the broader national experience with racial segregation. Both the 1917 U.S. Supreme Court decision in Buchanan v. Warley, as well as the 2008 decision in Parents Involved in Community Schools v. Seattle School District No. 1, were centered on practices of state-sponsored racial segregation in Louisville.
along lines of race and class (Louisville Metro Human Relations Commission, 2014), understood in the collective geographical imagination through the lens of the so-called “9th Street Divide”, a colloquialism referring to the traditional boundary between the city’s predominantly poor and black West End neighborhoods and the central business district.

Since the election of Mayor Greg Fischer in 2010, the West End has received an increasing amount of attention from the municipal government, with one Metro councilman from the almost entirely white and affluent far East End suburbs even declaring that “the 300,000 people east of Bowman Field do not exist in terms of what the mayor thinks about” (Bailey, 2014). From issues of vacant and abandoned property to air quality, the city has begun investing more time and money in addressing the range of challenges faced by these neighborhoods, in some ways spurred by a recent cover story in Louisville Magazine documenting the staggering inequalities between the West End and the rest of the city (Crutcher, 2013).

A key element of the collective social imagination of Louisvillians, as noted in both the Louisville Magazine story and in other venues, is the notion that the West End is fundamentally separate and apart from the rest of the city, with the aforementioned problems being isolated within these areas; the rest of the city does not have to deal with the ill effects of these issues, nor did they help to create them. The conceptualization of an isolated West End is reinforced by discussions of intra-urban mobility more generally. For example, in a presentation to the Louisville African American Initiative’s West Louisville Economic Development and Housing Summit on July 11, 2014, Fischer argued that mobility is one of the most persistent problems in Louisville, leading people of all ages to remain segregated; people from the East End tend to stay in the East End, people from the West End stay in the West End, people from the South End stay in the South End.

Because such understandings of the West End and its relationship to the rest of the city are clearly at conceptual odds with the relational understanding of space we use in this paper, we use the rest of the paper to empirically document that the conventional wisdom around socio-spatial segregation in Louisville is overly simplistic. While we do not discount the importance of both public policy and individual actions in producing the uneven landscapes of segregation in the city today, we demonstrate that the West End, through its relations to the rest of the city and the everyday mobilities of the people who live there, is not defined by segregation alone, nor is this process of segregation total. It is rather a complex, partial and selective process that affects some people and places more intensely than others.

4.1. Everyday mobilities and activity spaces

The first step in our analysis is analyzing the relative patterns of mobility of each of our two groups: West End and East End Twitter users. To do so, we calculated the odds ratio of each of the hexagonal cells spanning the greater Louisville area, using the adjusted tweet count, in order to show the relative presence or absence of each of the two groups throughout the city (see Fig. 3). The odds ratio indicates the prevalence of one group relative to the other in a particular place, while adjusting for the overall amount of tweeting from each of the two groups in that place. Values approaching 1 indicate that there is relative parity, while values approaching 0 showing a much greater prevalence of West End tweeting and values greater than 1 showing a greater prevalence of East End tweeting.

![Fig. 3. Unevenly segregated activity spaces of West End and East End residents.](http://dx.doi.org/10.1016/j.landurbplan.2015.02.020)
As expected, the areas used as the bounding boxes for classification of the two neighborhoods both show a general dominance of those users. But more useful insights can be gained by looking at the extent to which locations outside of these areas demonstrate high levels of either heterogeneity or homogeneity in tweeting activity and, thus, physical presence of West End and East End users. At first glance, the 9th Street Divide is clearly evident as the boundary at which a significant dominance of West End users gives way to greater levels of heterogeneity in the downtown area. But since the idea of the 9th Street Divide is generally used to denote the ‘failings’, so to speak, of the West End, including the segregation and constrained mobility of its residents, it is important to note that it is actually East End users who are more spatially segregated and confined in their everyday activity spaces, which will be discussed more below.

Given the large number of places with a relative prevalence of West End tweeting that are quite distant from the West End, it is evident that West End Twitter users are relatively mobile within the Metro area, traveling across the Ohio River to southern Indiana, or to neighborhoods like Hikes Point, on the periphery of the East End as we have defined it here. Indeed, only focusing on hexagons with odds ratios less than 0.5 as indicators of West End users’ mobility actually tends to underestimate the prevalence of West End users in other parts of the city (compare areas in purple to concentrations seen in Fig. 2b). Traces of this spatial mobility are also evident in isolated areas around major travel routes, such as I-64 or I-264, in the East End, where there is a lower relative amount of East End tweeting than in the surrounding residential areas. This initial look is interesting insofar as it offers a counterpart to the conventional narrative that class and racial segregation in housing is mirrored by limited mobility by these groups (cf. Cass, Shove, & Urry, 2005), especially in a city without significant alternative transportation options such as Louisville, though of course such mobility by West End residents is likely driven by the necessity of seeking out employment and consumption activities outside of their home neighborhoods.

In contrast, residents of the predominantly white and relatively affluent East End exhibit constrained activity spaces in that there are many parts of the city, especially the West End, to which they do not travel. For example, 45% of the adjusted tweets from West End users were east of 9th Street, while only 7% of tweets from East End users were west of 9th Street (see Fig. 4). This contrast is all the more stark when accounting for the fact that most of the South End neighborhoods, which demonstrate the most significant connection with the West End, are also west of the 9th Street boundary. The lack of presence of East End Twitter users is particularly intriguing given that the West End has become an increasingly important object of attention for policy elites (predominately from the East End) in the city. From a film and discussion series on urban renewal put on by a local historical society to a workshop about the rehabilitation of vacant properties in the West End, there are numerous examples of the public discourse around the West End that are dominated by individuals or groups with little or no direct connection to these neighborhoods. Indeed, it was the influence (or lack thereof) of outsiders that was seen a primary determinant in Metro Councilwoman Attica Scott’s failed reelection bid in the 2014 primary elections.

Despite this broader interest in the current and future state of the West End, it is evident that many East End Twitter users rarely come into actual contact with the West End, seemingly avoiding these neighborhoods almost in their entirety. This is not, of course, particularly surprising. As one local journalist summarized the relationship between the West End and East End:

Any Louisvillian who has lived here for more than a few years knows, almost instinctively, the boundary line between west

Louisville and the rest of Louisville: Ninth Street. Most white Louisvillians know it because they’ve heard some variation of the warning, ‘Don’t go west of Ninth Street.’ . . . Although the notion that west Louisville is a dangerous and even foreign place is embedded in the mental map that many of us – even the most bleeding-heartedly liberal and racially tolerant, if we will admit it – carry around in our heads, it is rarely talked about in public (Crutcher, 2013:25).

Indeed, this is even reflected in more popular expressions, such as the closing line of one local’s take on a popular meme, entitled “Shit Louisville People Don’t Say”, in which the white male narrator says ironically from the seat of his car, “Hey, do you want to go down to the West End?”

Thus, despite the salience of the 9th Street Divide metaphor, and its broader importance in drawing attention to socio-spatial inequality in Louisville, it clearly belies the complexity of how these inequalities are manifest in the everyday activity spaces of Louisvillians. When local real estate developer and civic leader Gill Holland

---

2 Video available from http://youtu.be/7Ru1qNDSP-o.
called 9th Street “the Berlin Wall of our community”, he likely meant to reinforce the notion that it is a rigid boundary. But like the Berlin Wall, the 9th Street Divide is not insurmountable, though movement across such divides were and are largely unidirectional, encouraging some to cross, while others remain prohibited or discouraged from such movements based on their social positions and broader structural forces. West End residents regularly travel outside of their home neighborhoods toward educational, employment and consumption opportunities further east and south, while many fewer East End residents move to their west when performing similar activities. These dynamics suggest the need for a more nuanced understanding of the 9th Street Divide, not as a border with fixed, defined and unchangeable characteristics that physically divides the city, but as a kind of spatial imaginary that is not entirely mirrored in people’s material spatial practices.

4.2. Exploring the fluidity of neighborhood boundaries

Building upon our analysis of the activity spaces of West and East End Twitter users, we now turn toward a broader reinterpretation of the neighborhoods and their boundaries based on the everyday mobilities of residents. Fig. 5 provides a simple visualization of those hexagons with 50 or more adjusted tweets for each user group, so as to provide a more expansive and spatially extensive definition of these neighborhoods than is possible with the use of conventional census tracts or area units. Our redrawing of these boundaries points toward the fluidity and porousness of the neighborhoods; while the West End and East End remain spatially distant enclaves in some respects, they also overlap at key points, such as the downtown and waterfront area, as well as suburban malls.

Our redefinition of the East End is rather subtle, largely a result of the fact that many more tweets from East End users fall within our original boundaries than is the case for West End users (compare Figs. 1 and 2c). But East End users, by and large, tend to gravitate eastwards, toward the outer suburban areas of Hurstbourne and Middletown, and commercial areas like the upscale mall The Summit. In addition, this redefined spatiality of the East End encompass much of the city’s downtown area as well as the traditionally working-class white neighborhoods of Germantown and Schnitzelburg, suggestive of recent urban redevelopment seeking to draw people back to the city’s downtown. The only point at which our expansive East End boundary approaches the conventionally defined West End is in the areas surrounding the Churchill Downs racetrack and the University of Louisville (see Fig. 7).

The redrawing of the West End is, however, much more significant in scope. While the census tract definition is bounded by 9th Street to the east, Algonquin Parkway to the south, and the Ohio River to the north and west, the everyday mobilities of West End Twitter users extend throughout the city. From the entire downtown area and the University of Louisville campus, to major transportation and commercial corridors throughout the East End, our redefinition of the West End demonstrates the incredibly partial story told through tropes of the West End as being somehow isolated and apart from the rest of the city.

For example, the activity spaces of the West End highlight a strong connection to the South End neighborhoods, a predominantly white and working class community sometimes characterized as the ‘redneck’ part of town in same way the West End is pejoratively labeled as ‘the ghetto’. For a city with such stark racial differences, it seems counterintuitive that the South

End would demonstrate a more significant connection to the West End through such everyday mobilities than with the similarly white and suburban areas of the East End. But, we would argue, this points toward the importance of combining such analyses of big data with situated, place-based knowledge that allow for an explanation of such anomalies, even if such an explanation is not definitive. In this case, the connection traces back to the movement of white working class families from the West End to the South End following the Great Flood of 1937, and the resulting influx of black families into these neighborhoods after WWII, itself a key moment of racial conflict in the city (Welch, 2013). Extending the metaphor of this kind of unstructured information as ’data exhaust’, we see the extension of the West End southward in Fig. 5a as the digital contrails of white flight, which continue to shape patterns of mobility within the city.

4.3. Scale dependent understandings in space and time

Thus far, we have largely highlighted those areas that represent differences in the spatial patterns and practices of geotagged tweeting by West End and East End users. We now turn to exploring in more depth those places that represent greater social heterogeneity in the city, as defined by greater parity in the levels of West End and East End tweeting in a given locality. But just as conventional understandings of the 9th Street Divide belie the complexity of relations between West End residents and the rest of the city, so too can our earlier analysis and methods of classification disguise what are still highly fractured social spaces as areas of heterogeneity and social mixing. While we can identify a number of relatively heterogeneous areas throughout the city (see areas in Fig. 3 represented by the hatched pattern for location quotient values between 0.5 and 2.0), this understanding is largely shaped by the methods we have utilized. By mobilizing alternative scalar framings (cf. Feick & Robertson, 2014), we can see that while West End and East End users are often incredibly spatially proximate to one another, they are rarely using the same physical spaces, and are even more unlikely to be co-present in the same places at the same times.

Using a finer scale analysis of one area of the city defined by relative heterogeneity, Fig. 6 demonstrates the significant ‘splintering’ effect made visible by this kind of social media data and a multi-scalar analysis (Graham & Marvin, 2001). In our initial analysis in Fig. 3, this area is characterized by varying levels of West End and East End tweeting, with some hexagons displaying relative heterogeneity, while others demonstrate a strong preference for one group over another. This is perhaps best illustrated in the block of businesses along South Preston Street (shown toward the upper left corner of Fig. 6), including Zanzabar, The New Vintage and Sunergos Coffee Shop which are all dominated by East End tweeting, consistent with their business models which cater to a younger, ‘hipster’ demographic in the neighborhood. In contrast, tweets from West End users in the area are much less concentrated, and tend to be located off of main thoroughfares in residential areas. Similarly, the area surrounding Audubon Hospital and a small shopping area (seen in the right side of Fig. 6) shows that at this scale, there are a couple of smaller areas of heterogeneity, surrounded by significant concentrations of West End tweeting. In addition, at this scale one can clearly see how the scalar shift influences measures of heterogeneity, as Audubon Hospital sits near the boundary of two hexagonal cells and is proximate to a largely wealthy and white, all-boys Catholic high school, with a dense concentration of East End tweeting, but which also straddles two hexagonal areas.

Such multi-scalar patterns of splintering urbanism are also evident at different temporal scales. Like the above analysis, some places that appear to be heterogeneous are actually marked by different uses across time, which can also impact our understanding of what’s actually happening on the ground in these locales. One such place is Churchill Downs, the historic horseracing track located just

3 For example, see the map of Louisville from the blog Judgmental Maps, available from http://judgmentalmaps.com/post/83423132066/louisville.
Fig. 5. (a and b) Redefining the boundaries of the West End and the East End.
to the southeast of the traditional West End boundary. In Fig. 3, Churchill Downs stands out as the lone cluster of East End tweeting west of Interstate 65 and inside of the Interstate 264 loop, and is surrounded by hexagons displaying varying levels of West End tweeting. But since much of the East End user presence in this area is related to horseracing, itself a seasonal activity, this locale offers the potential to demonstrate the temporal splintering in activity spaces across our two user groups. While it is broadly evident that much of the East End tweets in the area are in and around the track, with West End tweets more dispersed, we subdivided tweets in this area based on those occurring in and around racing season (April to June and September to November), and those in other months when the track is not active.

Fig. 7 uses a finer spatial scale similar to that seen above in Fig. 6 to show the relative amount of tweeting between West End and East End users during racing season (Fig. 7a) and outside of racing season (Fig. 7b). While there are many more West End tweets (n = 4191) than East End tweets (n = 1355) in this area, both user groups tend to tweet more from these areas during racing season, and at nearly equivalent rates. It is evident, however, that the places these users tweet from during the two time periods are significantly different; East End users tweet largely from Churchill Downs and the nearby Papa John's Cardinal Stadium during the months of racing season, while these concentrations diminish in non-racing months. West End users tend not to tweet from these places in any significant amounts, with tweets distributed throughout the surrounding neighborhood. We can thus see the dynamism of the area around Churchill Downs through time: most often this area is a spatial extension of the West End, though it temporarily transforms into a space of (relatively) elite consumption and a site of global attention, despite still being surrounded by a neighborhood from which many visitors have little connection whatsoever.

4.4. Contextualizing data practices through qualitative analysis

Up to this point, much of our analysis has followed in the tradition of research using mobile phone records or GPS trackers to study everyday mobilities and segregation. Apart from the fact that the use of geotagged social media data is considerably more cost-effective and less invasive for longer-term studies of mobility, one of the biggest advantages of using the massive databases of social media data available to us, however, is that it allows for greater attention to social context, rather than simply providing a record of presence in a given location at a given time. With geotagged Twitter data, we have some insight into what a given person was thinking or talking about, how they describe themselves, and with whom they were communicating, among other things. The qualitative content embedded in each individual tweet thus provides another means by which we can understand how Louisvillians move through, inhabit and experience the city.

To further target our qualitative analysis of tweets, we collated all tweets our user groups that fell within the boundaries of hexagons we determined to be ‘highly heterogeneous’—defined by a location quotient between 0.8 and 1.25 and a total of 100
Fig. 7. (a and b) Temporal differences in tweeting near the Churchill Downs Race Track.
tweets from West End and East End users combined—resulting in totals of 1812 and 1690 tweets, respectively. Hoping to further explore experiences of heterogeneity in places around the city, we conducted an inductive and iterative analysis of how users were engaging with these places, while keeping in mind that there is no necessary relationship between the location in which a user produces a tweet and the content of the tweet itself.

The first pattern that emerged from this analysis was the significant difference in the ways users’ tweets were connected to Foursquare check-ins, an often associated geosocial media platform focused on location-based social networking. In this sample of tweets in highly heterogeneous areas, 10.7% of all East End user tweets (n = 193) were associated with Foursquare check-ins, from a total of 44 separate users. On the other hand, there were only four Foursquare check-ins from four separate West End users, representing just 0.2% of all West End user tweets in these areas. This disparity, combined with the socio-spatial splintering noted earlier in Figs. 6 and 7, suggests that the socio-spatial practices and imaginaries of these two user groups are fairly disparate, even when they are operating in physically proximate areas.

While somewhat speculative, we would suggest that the relatively high Foursquare activity by East End users is consistent with theories of ‘conspicuous mobility’, or “[the] re-figuring of everyday mobility as a consumptive activity” (Wilson, 2012:1271) based on the sharing of (particular kinds of) locational information. In this case, the check-ins by East End users highlight their presence at ‘hip’ places – coffeeshops, bakeries, restaurants, bars and music or other entertainment venues – and serve to make these mobilities and consumption practices more known to others. In contrast, West End users active in these heterogeneous areas demonstrate little of this kind of effort toward locational visibility and instead were more likely to engage in distanced social interactions with individuals who are (seemingly) not co-present at that particular location, providing further weight to our understanding of the West End as being spatially diffuse and more appropriately defined by the density of social connections that stretch across urban space.

Another pattern we were able to identify was the relatively frequent references to ‘the ghetto’ and other associated terms amongst both East End and West End users. We chose to return to our original sample of East End and West End user tweets and map the spatial distribution of references to ‘ghetto’ in our entire corpus of tweets, which yielded 197 tweets from West End users and 87 from East End users. Interestingly enough, the spatial signature of the ‘ghetto’ in geotagged tweets amongst the two groups roughly mirrors the broader spatial patterns of these two user groups. There are very few East End tweets west of 9th Street, and while West End tweets are generally more concentrated within the conventional boundaries of the West End, there are also plenty of tweets outside of this area as well. Fig. 8 shows the distribution of these tweets, with West End tweets in purple and East End tweets in orange, with the text of selected tweets from East End users referencing the ‘ghetto’ also included. Looking closer to this content, we can see that only one East End user tweeted about the ghetto from the West End, saying “Holy cow we are in the ghetto”, while another user at Ballard High School in the East End declared “Ballard is so ghetto”. Especially in the case of these selected tweets, there is a level of cognitive dissonance at play, with many of these tweets from East End users being produced in
predominantly white and affluent, and mostly suburban, areas, far distanced both socially and spatially from anything that might resemble 'the ghetto'. Ultimately, these kinds of incongruencies demonstrate the more complex relationship between urban spatial imaginaries and the everyday activity spaces of individuals and collectives as demonstrated through geotagged social media data.

5. Conclusion

In this paper, we have developed a conceptual and methodological approach to the study of geotagged social media data that responds to earlier calls to go ‘beyond the geotag’. Rather than simply plotting the locations of individual twiters on a map, our approach combines relational socio-spatial theory with a variety of methods drawn from critical GLScience in order to place individual data points in relation to one another and to their broader social and spatial contexts through a more deliberate process of data collection and analysis.

In our case study, we focused on issues of intra-neighborhood segregation, mobility and inequality in Louisville, Kentucky, highlighting the fundamentally fluid, networked and relational nature of places in the city, as well as the dynamism of how people live in and occupy these places. Our analysis provides a strong counter-argument to the pervasive socio-spatial imaginary within the city of a ‘9th Street Divide’ that tends to isolate and pathologize the West End and its residents. But by understanding how people from different parts of the city actually move through and experience the city differently, we are able to demonstrate the contours of a more complicated set of socio-spatial mobilities that define the city and its neighborhoods through their extralocal relationships to other people and places. While this work has focused in particular on rethinking the socio-spatial imaginaries connected to particular classed and racialized neighborhoods in Louisville, future research with this kind of data could just as well focus more explicitly on issues of age, gender, sexuality or other identities as inferred from user profiles. Arguably, through classifying individuals based on these demographies is challenging and beyond the scope of this paper, it represents an opportunity to build upon our understandings of how this data can reveal a more complex and nuanced set of socio-spatial relations than is typically assumed.

Ultimately, we wish to reiterate that we are not arguing that geotagged social media data is an unequivocal improvement on, or replacement for, other forms of social and spatial data, especially when analyzing questions of inequality. But rather than reinscribing these inequalities through the use of such datasets, we would argue that our analysis has shown that this kind of social media data represents a potentially rich source from which to construct empirically-grounded counter narratives of these inequalities and popular socio-spatial imaginaries thereof, which in turn can allow for alternative conceptualizations of, and into actions into, urban socio-spatial relations and processes.

References


Barnes, T. J. (2009). ‘Not only...but also’: Quantitative and critical geography. The Professional Geographer, 61(3), 292–300.


Field, K. (June, 2014). I’m wondering when people will really realise the animated ecotoplasm twitter maps don’t actually show anything (http://t.co/sVvYlBn1f). In Tweet. Available from https://twitter.com/kennethfield/status/487877551038674128.


Poorthuis, A. (Place in New York City and Amsterdam: Exploring the possibilities and limitations of Volunteered Geographic Information. Amsterdam: University of Amsterdam (MSc. Thesis).


Taylor Shelton is currently a Ph.D. candidate in the Graduate School of Geography at Clark University. His research interests lie at the intersection of urban socio-spatial theory, internet geography and critical GIS, with a particular interest in new conceptual and methodological approaches to the study of geotagged social media data.

Ate Poorthuis is currently a Ph.D. candidate in the Department of Geography at the University of Kentucky. His research interests lie at the intersection of urban and internet geographies, with a particular focus on using mixed methodologies to study the everyday use of public space.

Matthew Zook is currently Professor of Geography at the University of Kentucky. His research interests lie in both economic and information geographies, with an interest in the role of code and new digital technologies in the shaping of everyday life.