Journal of Urban Technology

ISSN: 1063-0732 (Print) 1466-1853 (Online) Journal homepage: www.tandfonline.com/journals/cjut20

Revisiting Image of the City in Cyberspace: Analysis of Spatial Twitter Messages During a Special Event

Junfeng Jiao, Michael Holmes & Greg P. Griffin

To cite this article: Junfeng Jiao, Michael Holmes & Greg P. Griffin (2018) Revisiting Image of the City in Cyberspace: Analysis of Spatial Twitter Messages During a Special Event, Journal of Urban Technology, 25:3, 65-82, DOI: 10.1080/10630732.2017.1348881

To link to this article: https://doi.org/10.1080/10630732.2017.1348881

JOURNAL OF URBAN TECHNOLOGY
2018, VOL. 25, NO. 3, 65–82
https://doi.org/10.1080/10630732.2017.1348881

Routledge
Taylor & Francis Group

Revisiting Image of the City in Cyberspace: Analysis of Spatial Twitter Messages During a Special Event

Junfeng Jiao[a], Michael Holmes[b] and Greg P. Griffin[c]

[a]School of Architecture, The University of Texas at Austin, Texas, USA; [b]Center for Media Design, Ball State University, Indiana, USA; [c]School of Architecture, The University of Texas at Austin, Texas, USA

ABSTRACT

This research investigated people's communication of urban space as reflected in Twitter messages (tweets) during the 2012 Super Bowl. The authors archived over 600,000 tweets related to the Super Bowl from January 23 through early February 6. The authors identified 78 Indianapolis-area places or routes named in the tweets. Based on occurrence of these terms, the authors retained 9,103 city-specific messages for analysis. The frequency of such tweets changed over the two-week period and peaked two days before game day. Instances of all of Lynch's (1960) *The Image of the City* elements (node, district, landmark, path, and edge) were found in the tweets. While node-referencing terms were most common among the 78 spatial identifiers, district and landmark references were most common in the tweet sample. Edge references were almost non-existent and only occurred as named waterways. This research has implications for city-oriented social media monitoring efforts for future special events.

KEYWORDS

Urban space; Twitter; communication; Super Bowl; image of the city

> The Super Bowl's first-ever social media command center was an "enormous success," according to Taulbee Jackson, who managed the host committee's interactive communications hub. A team of strategists, analysts and tech-savvy volunteers spent the past two weeks monitoring the digital fan conversation while working out of a 2,800-square-foot space in downtown Indianapolis just blocks from where Super Bowl XLVI was played Sunday. They chimed in as needed via Twitter, Facebook, and other platforms … With some 150,000 people expected to flood downtown Indianapolis for Super Bowl festivities, the command center functioned as an innovative way to keep football fans informed and under control. (Laird, 2012)

Mental maps have been used to understand people's perception of cities and neighborhoods (Lynch, 1960; Long et al., 2007). These maps are typically operationalized as hand-drawn sketches intended to guide a visitor. When cities host major special events (e.g., conventions and sports events), locals and visitors face the challenge of communicating their understanding of urban space. Visitors tend to have shared goals and agendas (to attend particular events at a particular time) which drive spikes in demand for local resources (e.g., lodging, food, entertainment, transportation, and parking). With the fast growth of social network services (SNS, e.g., Facebook, Twitter, FourSquare), locals and

CONTACT Junfeng Jiao jjiao@austin.utexas.edu School of Architecture, The University of Texas at Austin, 310 Inner Campus Drive, Austin, TX 78712, USA

visitors increasingly tend to share their knowledge and questions about cities and neighborhoods through SNS postings. The Indianapolis Host Committee for the 2012 National Football League Super Bowl recognized the potential of SNS to enhance the visitor experience and arranged for a local firm, Raidious, to operate a social media monitoring center. This center, operated by company personnel and staffed by volunteers, monitored social media traffic for the two weeks up to and including game day (January 23 through February 5) and responded to questions and concerns about the Super Bowl host city with information about lodging, parking, traffic, transportation alternatives, and venue and event schedules. During the same two-week period, the authors archived Twitter traffic containing one or more of 12 Super Bowl and Indianapolis-related hashtags (words or phrases, prefaced with "#" used to define a tweet as belonging to a particular topic for purposes of search). Analyzing these postings for spatial references to places and paths within the Indianapolis metropolitan area reveals how urban space is communicated in the textual realm of SNS traffic, with implications for Lynch's (1960) long-standing framework for "the image of the city" and for future city-oriented social media monitoring.

Background

How people perceive urban spaces has been an interesting question for urban planners, policymakers, psychologists, and even marketing researchers for decades (Hospers, 2010; Lynch, 1960, 1984; Pearce and Fagence, 1996; Zmudzinska-Nowak, 2003). Kevin Lynch published his seminal book *The Image of the City* in 1960. Since then it has exerted enormous influence in urban design, environmental psychology, marketing, and social science disciplines. Lynch investigated 60 residents' mental maps or perceptions of their home cities (in his study, Boston, Los Angeles, and Jersey City). He found that a city's "imageability" and people's wayfinding process affected their perception of urban space. Lynch (1960) argued that urban space is imagined based on five different elements:

(1) Paths: Streets, sidewalks, trails, and other channels through which people move. Often paths are transportation related and serve as the backbone to connect different built environment structures.
(2) Edges: Boundaries or transition zones between two areas such as walls, buildings, and shorelines. Natural (river and mountain) or man-made barriers (highway and rail track) often become perceived edges.
(3) Districts: Large sections of the city distinguished by some identity or character such as neighborhood, downtown, or quarters.
(4) Nodes: Focal points, intersections or loci, such as squares, junctions, or stations.
(5) Landmarks: Easily identifiable objects which serve as public reference points, such as monuments and stadiums.

Further research showed not only local residents, but also visitors, commuters, and even people who have never visited the city also perceived cities based on these five urban image elements (Hospers, 2010; Knox and Marston, 2012; Pearce and Fagence, 1996).

Since its publication, the book has inspired numerous successors in a variety of disciplines, from Carr et al.'s (1984) research on how to design with city images in mind, Morello and Ratti's (2009) exploration of how to reinterpret Lynch's theory with 2D or 3D isovists, to Hospers' (2010) investigation of how to use Lynch's principles to shape city marketing, and Raynsforda's (2011) investigation of civic art and the aesthetic origins of Lynch's theory.

Other researchers have also investigated people's imagining/navigation behavior in new built environments. British sociologist John Urry argued the main reason for tourism and sightseeing is to look for visual experiences differing from what is normally seen at home. The dominant activity of tourists is "gazing" at "signs," which he describes as socially organized and systematized in a manner similar to how trained professions view important features in their work (Urry and Larsen, 2011). Crandall et al. (2009) found that visitors were more likely to take pictures at spots which had already been heavily photographed. The most recent studies leveraging SNS data in urban studies include use of geolocated media to analyze specific places. Researchers in Japan connected these new methods with more traditional observational analyses, finding that crowdsourced textual information could bridge behavior in the real world over large population sizes with the image of the city (Lee et al., 2013). These authors leveraged the latitude and longitude trace that can be extracted from Twitter's application program interface (API), and then also geolocated tweets using text information that could be located with a simple Google search. The study focused on the spatial behavior of the crowds, and noted future work should consider "exploiting much deeper crowd's minds and complicated crowd behavior by improving our approach with the analysis of textual messages of tweets … " (2013: 619). Communication researcher Mor Naaman (2011) had previously proposed that it would be possible, and useful, to classify social media data into Lynch's framework, but specific methods were needed to realize this goal.

However, it remains unclear how people—including visitors unfamiliar with a city—perceive and communicate about urban space in cyberspace and whether Lynch's (1960) theory can be applied to the textual realm of SNS traffic. Most recently, an international team of researchers proposed that

> Lynch's (1960) concept of mental maps of the city … can now be extended and explored on a massive scale through crowdsourcing. This requires not only deriving new knowledge from such data, but also exploring how such knowledge could be studied through the lens of Lynch's work (Crooks et al., 2015).

No research to date has explored whether people's city descriptions as communicated online show similar characteristics as demonstrated in previous research on cognitive maps of urban spaces.

Twitter traffic has proven a useful research resource in domains as varied and important to urban affairs as public health (Chew and Eysenbach, 2010), transportation planning (Evans-Cowley and Griffin, 2012) and emergency management (Hughes and Palen, 2009). Lynch's (1960) *The Image of the City* remains a valid approach to understanding the way people, including tourists, perceive space (Pearce and Fagence, 1996). We believe special events such as the 2012 Super Bowl present a rich context for application of Lynch's theory because such events bring large numbers of visitors to a city and generate large volumes of Twitter traffic. As resident and visitor needs, concerns

and movements within the city will vary with the calendar of events leading up to Super Bowl game day, and may also vary with daily patterns of activity, we incorporate change over time into our exploration of spatial references in Twitter traffic about the Super Bowl.

Research Questions

Three research questions shape this exploratory research. Each reflects our desire to investigate how urban spaces are described and inquired about in short-form SNS traffic in the two weeks up to and including Super Bowl Sunday, 2012:

RQ 1. What spatial words and phrases are used by people when they share city-related questions and answers via Twitter posts during a special event?

RQ 2. What is the distribution in Twitter traffic of event-related spatial references across Lynch's cognitive map elements of node, landmark, path, district, and edge?

RQ 3A. Do spatial references change in frequency and categorical distribution over the days leading up to a special event?

RQ 3B. Do spatial references change in frequency by hour of the day?

Method

This study applies a three-step content analysis process: (1) archiving Twitter traffic with Super Bowl-related hash tags; (2) filtering the archive to retain only tweets naming places or routes in the Indianapolis area; and (3) coding the content of retained tweets according to the five elements of Lynch's (1960) framework for characterizing "images of the city" (hereafter IC).

Archiving Twitter Traffic

A consumer-grade Twitter archiving tool, *The Archivist* (Microsoft, 2010), was used access the Twitter stream and record tagged tweets. Due to Twitter-imposed limitations on the volume of tweets allowed in response to any single search request, the 12 selected hashtags were distributed across four instances of the program, each running on a different computer (#SB, #SB2012, #SB46, #SUPERBOWL, #SUPER46, #SBXLVI, #XLVI, #SOCIAL46, #INDYSB, #INDY46, #INDY and #INDIANAPOLIS).

A complete and exhaustive record of Super Bowl-tagged tweets is theoretically possible but difficult to achieve. The constraints imposed by Twitter make it impossible to guarantee the archiving tool was able to return every tweet with a target hashtag within a time period (the software is limited to returning no more than 1,500 tweets every time the search is automatically refreshed). It is likely only a small fraction of total traffic was captured during the game itself, as Twitter volume exceeded 10,000 tweets per second at halftime (though what portion of these tweets contained our target hashtags is unknown). In addition, we identified official (publicized) hashtags and likely unofficial hashtags as targets for archiving. Other tags may have been adopted by users to label event-related posts and such tags may have risen and fallen in use over time. For example, if a celebrity visitor tweeted about being in the city for the game, people "following" that celebrity on

Twitter may have tagged event-related posts with that celebrity's hashtag but not with any of the common Super Bowl tags. However, we chose to limit the archiving to the same set of tags throughout the two-week period to support day-to-day comparisons.

Archiving began on January 21 to allow testing of the logging setup. Only tweets from January 23 to 9:00 A.M. on February 6 were retained for this analysis. Twitter users often apply multiple hashtags in a single tweet so many tweets were recorded by several instances of *The Archivist*. As each tweet is assigned a unique ID number by Twitter, duplicate posts were easily identified and deleted after the four archives were joined into a single file.

Filtering the Tweet Archive

The archiving software did not provide tweet geolocation information so an extensive filtering process was required to identify and retain only tweets containing spatial references within Indianapolis (and therefore likely originating within the metro area). The process was as follows:

(1) Duplicate tweets (as indicated by identical tweet ID numbers) were deleted.
(2) Simple retweets were deleted (i.e., posts with the retweet marker, "RT," within the first 10 characters of the tweet; these posts added little or no content to the original post and therefore constituted another form of data duplication).
(3) Tweets in languages other than English were deleted (primarily Spanish, with isolated examples in other languages). These were flagged by searching for common Spanish and German words. Other instances were manually deleted when encountered in the data set.
(4) Tweets returned by city hashtags (#Indy, #Indianapolis) were deleted if they did not also contain a Super Bowl hashtag or Super Bowl content. Deleted tweets were typically weather or news updates from local media and job postings by local employment services.
(5) Tweets were flagged for retention if they included unambiguous references to places within Indianapolis and Super Bowl event venues and routes (e.g., downtown, Super Bowl Village, stadium, convention center, airport, city interstates, key streets and avenues; etc.; see Appendix). Tweets containing only a reference to the entire city (Indianapolis, Indy, Naptown) were not flagged for retention.
(6) In an iterative process, the remaining tweet database was manually reviewed for content patterns that could be expressed as exclusion rules. These filters were then used to delete event-related posts that did not contain references to spaces or places within Indianapolis. Common patterns included food and snack recipes, plans for Super Bowl viewing parties, expressions of support or non-support for the competing teams, predictions about the game's outcomes, commentary on halftime show artists or performances, and references to the teams' home cities and home stadiums.
(7) Any spatial reference recognized by the second author (a former Indianapolis resident and long-term resident of central Indiana) as designating a place within the Indianapolis metropolitan area was added to the list of words or phrases used to flag tweets for retention.

(8) During the manual scanning process numerous exclude/retain decisions were made about individual tweets; in each case the authors considered whether or not that decision could be reliably expressed as filtering rule to be used for automatic flagging of the data.
(9) Steps six through eight were repeated until exhaustive; that is, the process halted when review of the tweets yielded no new rules for excluding or retaining multiple tweets and few individual tweets were selected for deletion.

Step 7 of the process generated 78 Indianapolis-specific spatial words or phrases (See Appendix, Table A1). These terms, including their variant forms or spellings, served to identify tweets to be retained in the sample. Examples of deleted and retained tweets are provided in Table 1.

Tweet Content Coding

The researchers applied a classification system for spatial references developed from the five categories of Lynch's (1960) typology of elements of the image of the city (IC; Table 2). Each of the city-specific spatial words and phrases identified during content filtering was assigned by the authors to one of the IC categories in Table 2 to allow machine coding of the tweet content through substring pattern-matching searches. Of the 78 terms, 30 (38.5 percent) were nodes, 18 (10.3 percent) referenced paths, 16 (20.5 percent) named landmarks, 11 (14.1 percent) referenced districts, and only three (3.8 percent) designated edges.

The coding syntax allowed any given tweet to be assigned multiple categories (e.g., node and path or district and landmark) but not for multiple instances of a given category in a single tweet. A "node" category assigned to a tweet, for example, means the post contains one or more node references.

Results

This is exploratory research driven by research questions rather than hypotheses. Our analysis emphasizes description and cautious interpretation rather than statistical

Table 1. Examples of deleted and retained tweets (user names omitted for clarity)

Examples of deleted tweets	Rationale
#Madonna look-A-like contest after #halftime Winner gets 520 in CASH&prizes! #SuperBowl	Super Bowl reference without spatial content.
@TwoWheeledBeard @JBaty83 @HoraceRawlins Patriots Vengeance Tour 2011–12. Next stop #SBXLVI in Indy.	No reference to anything within the city.
@GilletteStadium: Good morning back! Awake, Alive, Still excited!! #SBXLVI here we come!	Stadium reference is a team home stadium, not Lucas Oil Stadium in Indianapolis.
RT @itsyourboySham: Patriots are going to, without question, win the #SuperBowl	Simple retweet; also lacks spatial content.
Just walked downtown and they already getting ready for the super bowl #Indy	Includes references to a portion of the city and to the Super Bowl.
Is there a wait at kilroys? #social46 #superbowl2012	Names a local restaurant and has two event-related hashtags.
Check out #Indy downtown! RT @TheFieldhouse: Plenty of events in and around @TheFieldhouse in the coming weeks …	Includes a neighborhood reference and a venue reference (also an example of a retweet with added content).
XLVI Roman Numerals are up at Monument Circle. #superbowl #indianapolis	Contains a Super Bowl reference and a local landmark reference.

Table 2. IC coding categories

IC Element	Definition
Node	A center of activity "which the observer can enter, and which are intensive foci to and from which he [sic] is traveling" (Lynch, 1960: 72)
Landmark	Prominent visual feature; for Lynch, typically not entered but observable. For the purpose of this study we include in this category large buildings and structures with visual "referenceability" when providing directions.
Path	"Major and minor routes of circulation" (47)
Edge	Dividing lines between districts; "linear elements not used or considered as paths" (62)
District	"Medium to large sections of the city … recognized as having some common identifying character" (67)

inference testing, in recognition of the preliminary nature of the study and the limitations on the completeness of the Twitter sample described earlier.

Profile of the Raw Sample

From 12:01 A.M., January 23, 2012 to 9:00 A.M., February 6, 2012, *The Archivist* logged 648,893 tweets containing one or more of the target hashtags. Deletion of duplicates from among the four separate *The Archivist* logs reduced the raw sample to 499,756 tweets. The average number of tweets per day was 33,317 with a SD of 65,812.69. A negatively skewed distribution of traffic over the two-week period is the source of the huge variance. Logged Twitter traffic was light for most of the study period with volume trending upward on Thursday, February 2 (See Figure 1), possibly due to an influx of visitors driven by the requirement by many downtown hotels that Super Bowl visitors reserve rooms for three or four nights. The peak logged daily traffic of almost 250,000 tweets occurred, unsurprisingly, on game day (Sunday, February 5).

The average number of tweets logged per hour was 1,448.6 (SD 3129.2). The average hourly traffic was over 1,000 tweets per hour from 8:00 A.M. to midnight (m=1907.1, SD=288.7) but dropped off sharply from midnight to 8:00 A.M. (m=413.7; SD=150.6; See Figure 2).

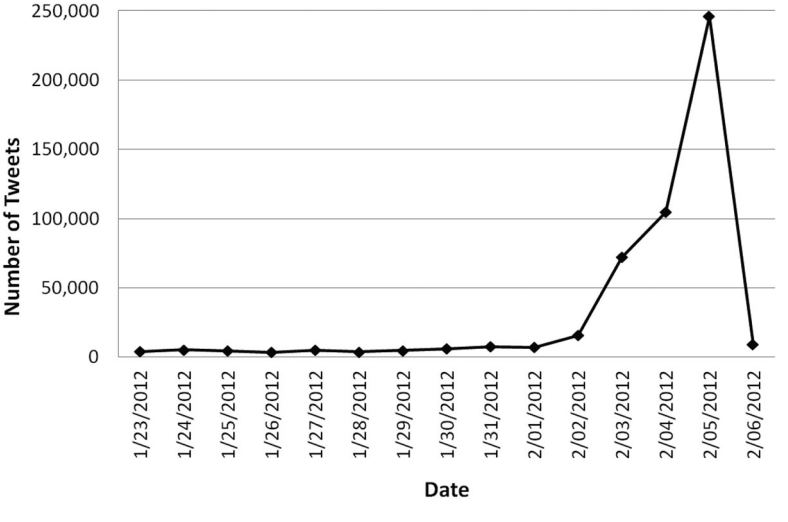

Figure 1. Distribution of tweets per day during the observation period

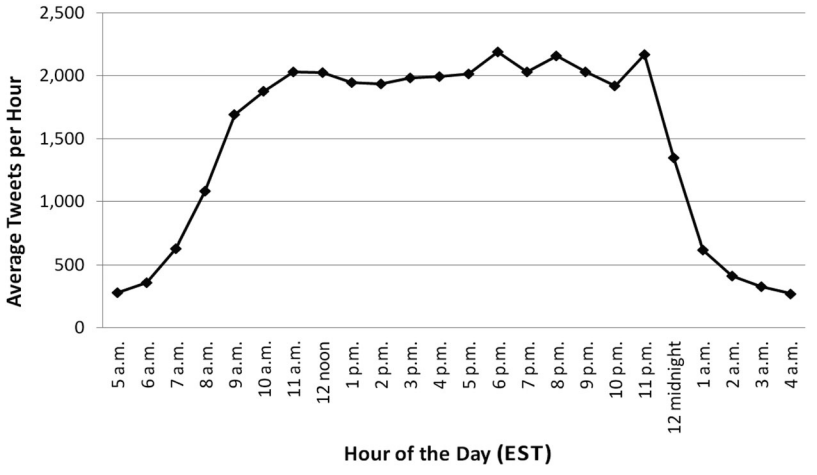

Figure 2. Distribution of average tweets per hour during the observation period

Profile of the "Spatial Tweet" Sample

From the nearly 500,000 tweets described above, 9,103 (1.8 percent) survived the filtering process to be retained as "spatial tweets" naming Indianapolis places and/or routes listed in the Appendix. Of these, 8,558 (94 percent) were original posts and 545 (6 percent) were retweets to which a user had added content. The average number of spatial tweets per day was 606.9 (SD 610.2). The peak volume was slightly over 2,000 spatial tweets on February 3rd, two days before Super Bowl game day (See Figure 3). The average number of spatial tweets per hour was 28.5 (SD 34.2). The hourly traffic takes on an approximately normal distribution over a 24-hour period (See Figure 4). This contrast to the hourly distribution for the raw data may be an artifact of the filtering process. Spatial tweets as defined here are more likely to originate within the Indianapolis area and therefore reflect time-based

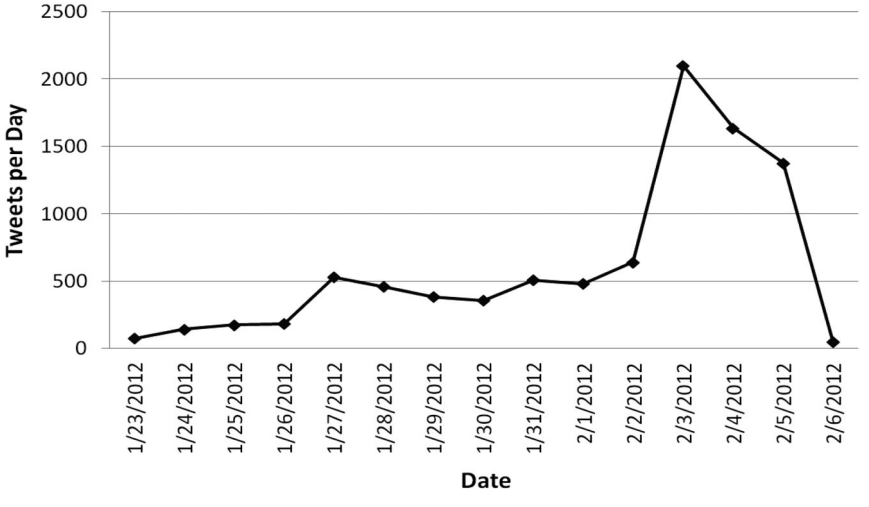

Figure 3. Distribution of spatial tweets per day during the logging period

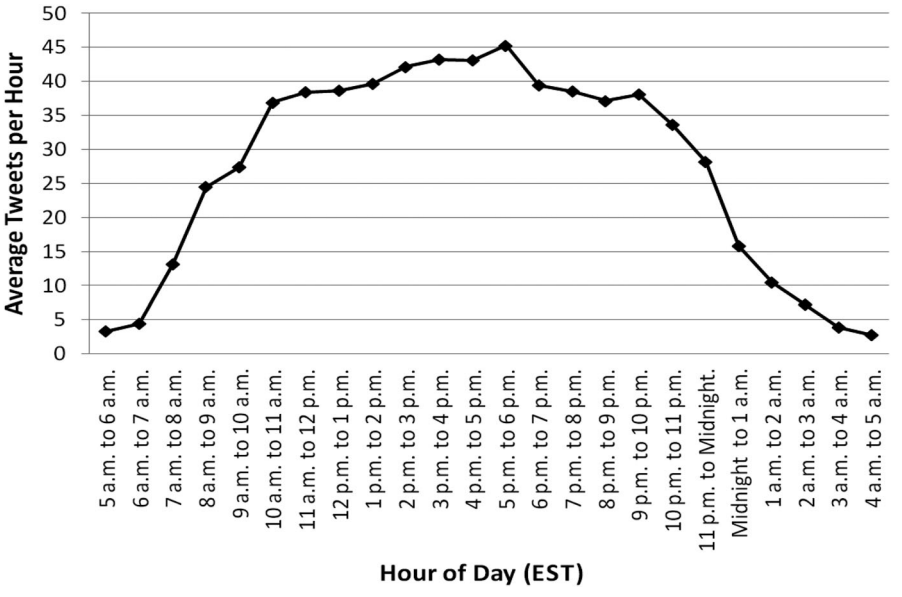

Figure 4. Distribution of average spatial tweets per hour during the logging period

activity patterns within a single time zone (Eastern). The raw sample, in contrast, doesn't have as discernable a late-afternoon peak because activity levels are smoothed by time zone differences. The 8:00 P.M. to 9:00 P.M. value in Figure 2 includes activity from 7:00 P.M. to 8:00 P.M. CST, 6:00 P.M. to 7:00 P.M. MST, etc., whereas we can assume almost all of the 8:00 P.M. to 9:00 P.M. traffic in Figure 4 is generated by users in Eastern Standard Time.

Figure 5, which displays tweets per hour for each day of the logging period, reveals that the increase in spatial tweet traffic on Thursday, February 2, begins in the afternoon. This finding provides some support for our conjecture that visitor arrival in the city on Thursday influenced city-specific Twitter traffic.

IC Category Results

As noted above, individual tweets could contain more than one type of spatial reference. In our sample 14.3 percent of the tweets yielded two IC categories and 1.2 percent yielded three categories. For example, a tweet from Indy2012Rentals, "NOTE: We are not accepting any new #SB46 listings outside Downtown #Indy, Broad Ripple and 2 miles from @LucasOilStadium" names two neighborhoods (districts) and one venue (landmark). Notice that implied districts not in local vernacular––in this case a circle of 3.2 kilometers (two miles) radius around the stadium––were not captured by the automated coding.

District references were most common, occurring in 64.8 percent of the spatial tweets, followed by references to landmarks (21.8 percent), nodes (19.5 percent), and paths (9.7 percent). Edge references were rare, with the 11 examples constituting just 0.12 percent of the sample and averaging less than one tweet per day.

Daily and hourly average frequencies for tweets referencing the IC categories are provided in Table 3. District was the most tweeted element, followed by landmark, node, path, and edge. The distribution of IC references varied over the two-week period

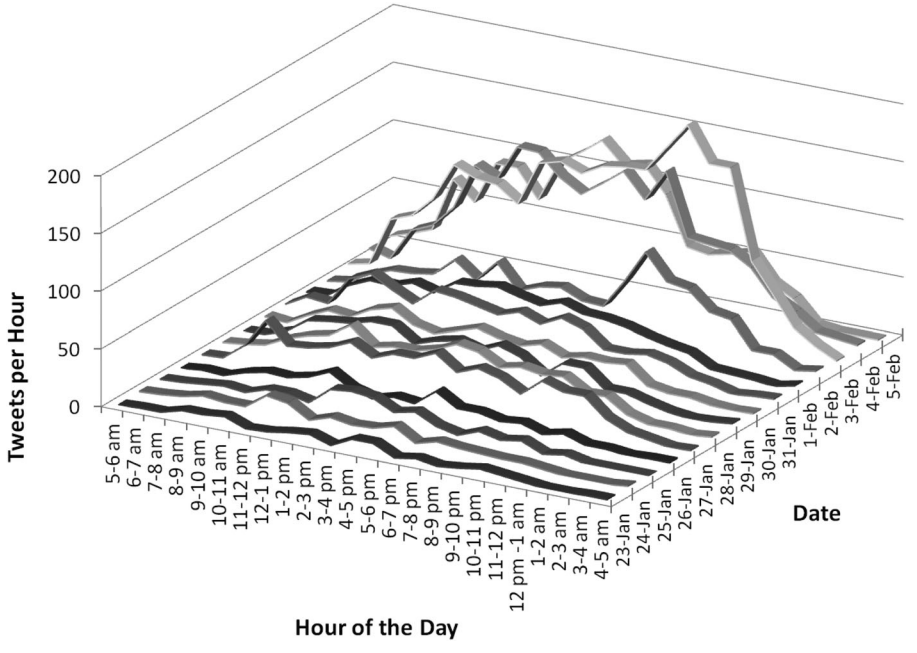

Figure 5. Tweets per hour by day (midnight to 5:00 A.M. for a given date are considered part of the previous day for this illustration)

(See Figure 6). District, node, and path references peaked on Friday, February 3, while landmark mentions peaked on game day, February 5, likely driven by stadium references.

The distribution curves over hours of the day are similar for the various elements (See Figure 7); however, there is an apparent, temporary early evening drop in mentions of districts and nodes concurrent with an increase in landmark references which warrants closer examination in future analyses.

While the number of spatial tweets per hour varied greatly through the day, the distribution of spatial references across the IC categories remained relatively consistent within each hour (See Figure 8). The only apparent trend in the distribution appears as a turning point around 7:00 P.M., when the share of node references diminishes slightly and the share of district references increases. This may reflect more attention to official evening events (concerts and other attractions) the downtown and village districts.

Table 3. Daily and hourly frequency of tweets containing IC category references

	Tweets per day*		Tweets per hour	
	Mean	SD	Mean	SD
Node	118.3	87.2	5.6	5.9
Landmark	132.6	181.5	6.2	10.8
Path	58.8	70.3	2.8	4.6
District	393.1	409.1	18.4	23.3
Edge	0.7	1.4	0.03	0.2

*A single tweet may contain multiple category references

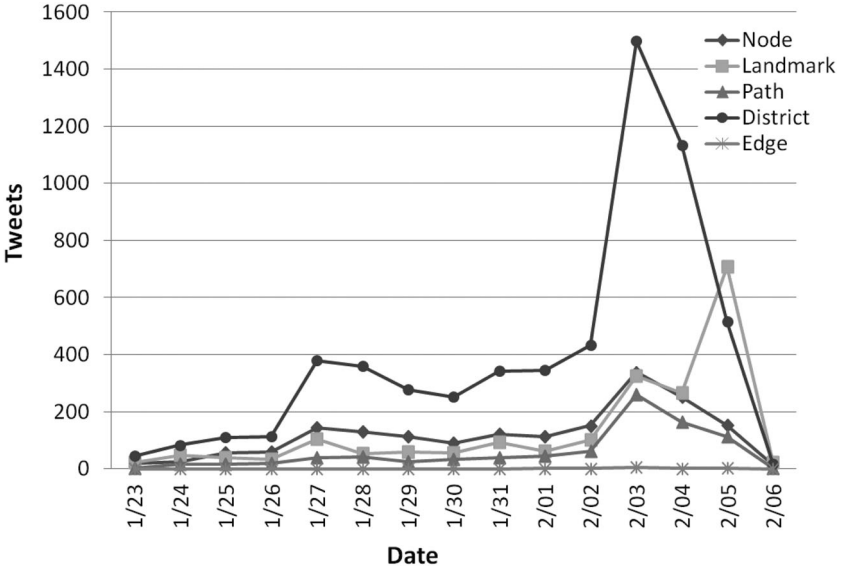

Figure 6. Tweets per day for the five IC spatial reference categories

Most Frequent Spatial Words

Two of the top 10 most frequently used spatial words in the sample were district elements (See Table 4). These places, downtown and village, were both foci of event-related activities. Landmark elements were also frequent in the form of stadium and Monument Circle

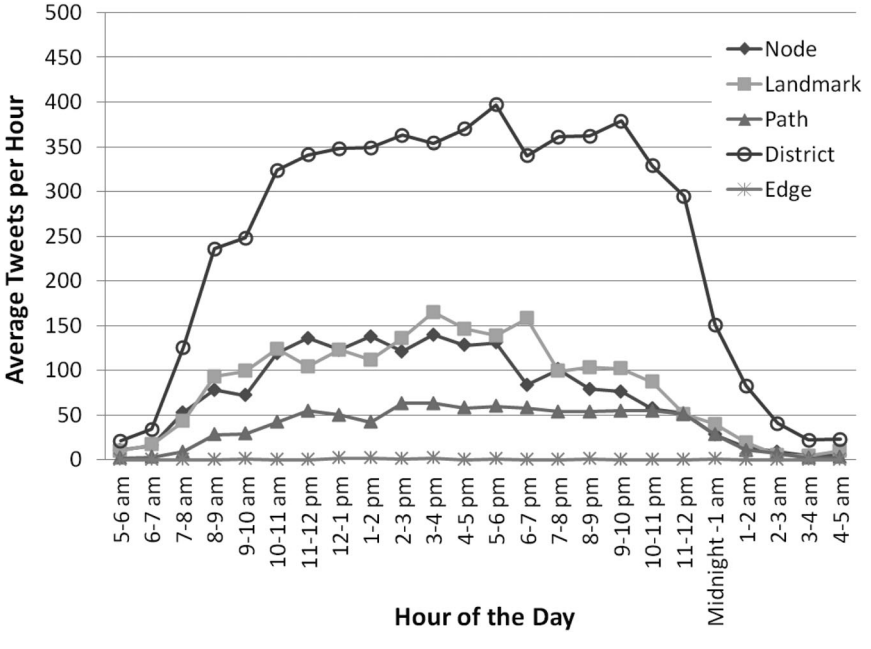

Figure 7. Average tweets per hour for the five IC spatial reference categories

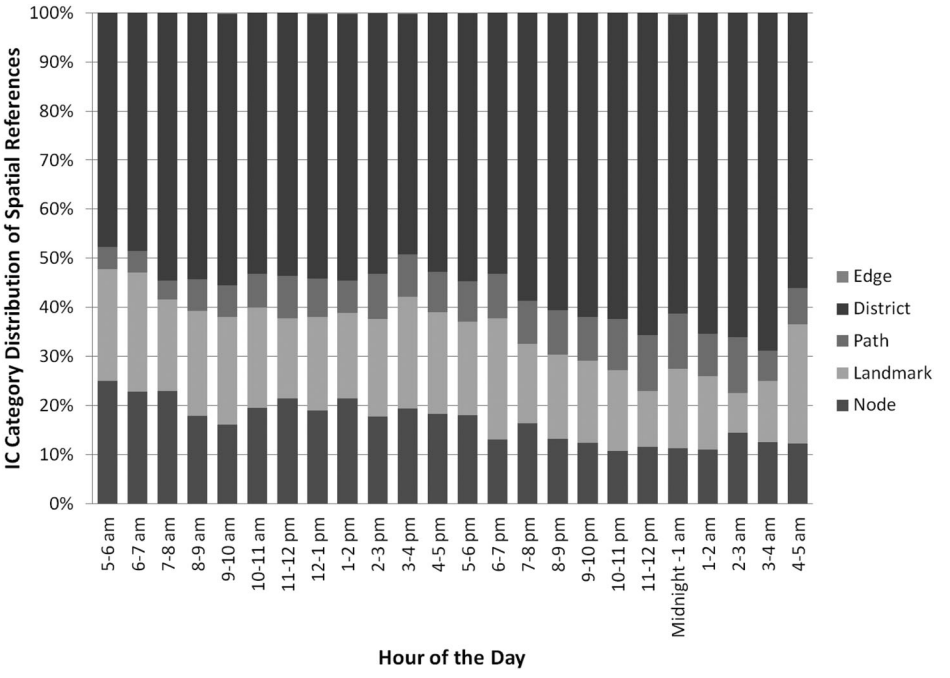

Figure 8. Distribution of spatial references per hour of the day

(the former was the game venue and the latter is the symbolic center of downtown Indianapolis). Three top 10 words were node references (zipline, hotel, and NFL Experience). Of these three, two were central entertainment venues during the run-up to game day. The remaining three words were path elements (Georgia, avenue, and street). Georgia's popularity is explained by the fact it was the pedestrian route through the Super Bowl Village. Avenue and street occurred frequently with specific route names (e.g., Meridian, Alabama, etc.).

Spatial terms can of course, be mapped. Lynch (1960) combined results of individual sketch maps and personal interviews into simplified maps of the "Image of the City" elements. Not all spatial tweets are necessarily mappable, however. Language varies in

Table 4. Most frequently occurring spatial terms and their co-occurrence with representatives of other IC categories.

				Frequency in Tweets Containing Other IC Elements				
Category	Rank	Total	Assigned	Node	Landmark	Path	District	Edge
1	Downtown	3261	District	293	191	193	--	4
2	Village	2791	District	305	124	131	--	1
3	Stadium	1218	Landmark	52	--	25	51	0
4	Georgia	768	(multiple)[1]	149	227	192	396	0
5	Avenue	754	(multiple)	158	122	199	447	1
6	Zipline	526	Node	--	12	11	196	1
7	Monument Circle	493	Landmark	80	--	32	133	0
8	Street	378	Path	32	26		142	0
9	Hotel	370	(multiple)	254	158	7	86	0
10	NFL Experience	362	Node	--	15	9	163	0

[1]Assigned category is dependent on semantic context and if used alone or in a phrase

Figure 9. Frequent spatial tweets mapped in the *Image of the City* typology

geographic specificity, so terms like "White River State Park" are easily identified, whereas a generic "avenue" could be any number of locations (Griffin, 2014). Figure 9 provides an example of Lynch's method adapted to discussion over social media about special events.

The map (See Figure 9) clearly shows that the most memorable IC element was the district (downtown and Super Bowl Village) followed by landmarks (Stadium, Marriott Hotel, Monument Cycle), nodes (Zip line, Tap room, Elbow room), and paths (Super route, Washington St, Meridian St). Edge was barely mentioned by people and only two waterways (White River and Canal) were identified.

Conclusions

Our first research question asked "What spatial words and phrases are used by people when they share city-related questions and answers via Twitter posts during a special event?" The results showed that among nearly a half million collected tweets, only 9,103 (1.8 percent) had spatial key words and could be classified as "spatial tweets," indicating that people were either unwilling, unable, or unconsciously forgot to include any spatial references in their tweets. Further iterative scanning and filtering processes yielded 78 Indianapolis-area places or routes named in the tweets. This might suggest that compared to the real world, our current virtual or cyber space is less spatially oriented

(Cao et al., 2015; Zhao et al., 2016). Although analyzing these spatial tweets can shed light on people's perception of urban spaces, the information is limited (Arribas-Bel et al., 2015).

The second research question addressed the distribution of event-related spatial references across Lynch's elements of node, landmark, path, district, and edge. All of the *Image of the City* elements were found in the Indianapolis-specific Twitter traffic. While node-referencing terms were most common among our 78 spatial identifiers, district and landmark references were more common in the tweet sample. Landmarks and nodes are usually distinct objects and easy to memorize, serving as reference locations for locales and helping visitors understand a new physical environment. They help human beings understand and remember large physical spaces, which, in turn, helps them navigate through these spaces (Caduff and Timpf, 2006). Research also shows that landmarks and nodes are important in people's wayfinding process and that people who were presented with more visual landmarks and nodes along routes had higher recognition scores and shorter response times than those who were not presented with them (Denis et al., 2014). Landmarks and nodes also play important roles in social media-based travel recommendations (Han and Lee, 2015) and, not surprisingly, were often included in spatial tweets. In addition to landmarks and nodes, districts, which describe a large area or a zone (e.g., downtown, Super Bowl Village), are also easy to remember and often become popular destinations for visitors. Thus, they were also frequently included in spatial tweets. Overall, these geo-markers served as important memory points, which helped visitors to understand and represent new physical spaces in SNS communications (Arribas-Bel et al., 2015; Caduff and Timpf, 2006; Zhao et al., 2016).

Unexpectedly, we found edge references to be almost non-existent and to occur only as named waterways; this may indicate their limited utility in short textual descriptions of urban space or a relative unimportance of edges in cognitive maps of a flat, geographically homogenous urban space such as Indianapolis where major district boundaries are typically paths and there are few other forms of physical barriers. This possibility suggests value in contrasting this text-driven analysis with analysis of a more traditional sample of hand-drawn maps of Indianapolis.

Our next concern was whether spatial references would change in frequency over the days leading up to and including Super Bowl game day. The overall frequency of event-related Twitter traffic and spatial tweets increased over time; however, they had a slightly different distribution pattern. Event-related traffic peaked on game day but spatial references peaked two days earlier, suggesting visitors and residents were concerned with different activities and intentions during the pre-game period and game day. The frequencies of different types of spatial references also differed by day. District, node, and path terms peaked on February 3, in contrast to landmarks which peaked on February 5. This raises the possibility that visitors were more likely to ask general location questions or describe general plans online when they first arrived in the city and had time to enjoy the many downtown and Village venues and official activities, with attention shifting to the stadium (a landmark) on game day.

The distributions over hours of the day for the five types of spatial references were not markedly different from each other, though district terms were most common in every hour. We conjecture the slight peaks in district references at 5:00 P.M. to 6:00 P.M. and

again four hours later might reflect declarations of plans or intent ("I'm going to the Village tonight") and subsequent reports ("Had a great time downtown!") but this remains a question for future exploration.

Implications, Limitations, and Future Directions

This research investigated people's perception of urban space as reflected in SNS communications during a special event. In addition to demonstrating the applicability of Lynch's (1960) framework to textual data, the research sheds light on use of social media to facilitate visitors' experiences of a new urban environment and their navigation through an unfamiliar space. For example, knowing that visitors will have many district, node, or landmark related questions may allow the host city of a major event (e.g., the Olympics, World Cup, World Expo, other major political, social, racing, and sporting occasions) to more efficiently identify useful search terms for social media monitoring. Anticipating the likely kinds of city-related questions and references can aid a host city's monitoring service in preparing stock answers and information resources. Being aware that host-city-related SNS traffic may peak before broader event-related traffic may help hosts organize their resources and adjust SNS monitoring efforts appropriately over time. Lastly, this paper presented an alternative way to generate Lynch's image of the city map. Compared to a traditional mental mapping method, researchers can easily collect a large amount of online SNS data and use it to generate more accurate city perception maps, which could be used for future large event planning and city management.

There are some limitations to this research. First, the message collection process was limited by the capabilities of the consumer-grade archiving software, which can only capture text-based content. Links to websites and video content were not recorded in this process. Second, the logging software was incapable of recording tweet geolocation information, which prevented comparison of the writer's location with the location referenced in the tweet. Third, some relevant tweets might have been missed due to our reliance on a predefined and unchanging set of hashtags. Fourth, the automated coding generated results with apparent face validity, but we have not yet assessed the filtering and coding processes for reliability. Finally, content analyses were not included in this paper. The authors did not explore why specific locations were mentioned in tweets or the intent of different tweets (Riff et al., 2014; Stellefson et al., 2014).

Special events are, paradoxically, common. There will be opportunities to redress the limitations of this research in the future by using a commercial-grade Twitter traffic archiving service. Such commercial services are not subject to the constraints imposed on *The Archivist* and will be able to provide a more exhaustive sample and geolocation data for tweet origination. The impact of systematic updating of hashtags and search terms can also be assessed. The reliability of the iterative filtering and automated categorization of messages can be tested by comparing results to manual coding by trained judges of random samples from the raw data.

Our analysis focuses solely on spatial language (places and routes) in the event-related Twitter traffic. Our next efforts will expand the analysis to explore the tweet text co-occurring with spatial references to discover if patterns of movement words ("going to," "headed for," "walk," "drive," etc.) and positive or negative evaluative terms emerge in consistent associations with the spatial terms.

Future research also presents the opportunity to broaden the scope of such SNS spatial content analysis to other issues of interest to urban studies scholars and professionals, such as traffic management, emergency response, perceptions of environmental quality, micro-climate studies, and smart city studies. This study not only confirms the applicability of Lynch's framework to special events, but the coherence of this spatial language over social networking services.

Disclosure Statement

No potential conflict of interest was reported by the authors.

Notes on Contributors

Junfeng Jiao is an assistant professor in Community and Regional Planning and director of the Urban Information Lab at The University of Texas at Austin. He led this Twitter Super Bowl Study.

Michael Holmes is a professor in Journalism at Ball State University. He worked on the research design, data analysis, and the composition of the paper.

Greg P. Griffin is a PhD student in Community and Regional Planning at the University of Texas at Austin. He assisted in the composition and revision of the paper.

Bibliography

D. Arribas-Bel, K. Kourtit, P. Nijkamp, and J. Steenbruggen, "Cyber Cities: Social Media as a Tool for Understanding Cities," *Applied Spatial Analysis and Policy* 8: 3 (2015) 231–247.

D. Caduff and S. Timpf, "A Framework for Assessing the Salience of Landmarks for Wayfinding Tasks," *Cognitive Processing* 1: 7 (2006) 23–23.

G. Cao, S. Wang, M. Hwang, A. Padmanabhan, Z. Zhang, and K. Soltani, "A Scalable Framework For Spatiotemporal Analysis of Location-Based Social Media Data," *Computers, Environment and Urban Systems* 51 (2015) 70–82.

S. Carr, L. Rodwin, and G. Hack, "Kevin Lynch: Designing the Image of the City," *Journal of the American Planning Association* 50: 4 (1984) 523

C. Chew and G. Eysenbach, "Pandemics in the Age of Twitter: Content Analysis of Tweets during the 2009 H1N1 Outbreak," *PLoS ONE* 5: 11 (2010) e14118.

D. Crandall, L. Backstrom, D. Huttenlocher, and J. Kleinberg, "Mapping the World's Photos," Proceedings of the 18th International World Wide Web Conference (Madrid, Spain, April 20–24, 2009).

A. Crooks, D. Pfoser, A. Jenkins, A. Croitoru, A. Stefanidis, D. Smith, S. Karagiorgou, A. Efentakis, and G. Lamprianidis, "Crowdsourcing Urban Form and Function," *International Journal of Geographical Information Science* (2015) 37–41.

M. Denis, C. Mores, D. Gras, V. Gyselinck, and M.-P. Daniel, "Is Memory for Routes Enhanced by an Environment's Richness in Visual Landmarks?" *Spatial Cognition & Computation* 14: 4 (2014) 284–305.

J. S. Evans-Cowley and G. Griffin, "Microparticipation with Social Media for Community Engagement in Transportation Planning," *Transportation Research Record: Journal of the Transportation Research Board* 2307 (2012) 90–98.

G. P. Griffin, "Geographic Specificity and Positionality of Public Input in Transportation: A Rural Transportation Planning Case from Central Texas," *Urban, Planning and Transport Research* 2: 1 (2014) 407–422.

J. Han and H. Lee, "Adaptive Landmark Recommendations For Travel Planning: Personalizing and Clustering Landmarks Using Geo-Tagged Social Media," *Pervasive and Mobile Computing* 18 (2015) 4–17.

G. J. Hospers, "Lynch's *The Image of the City* after 50 years: City Marketing Lessons from an Urban Planning Classic," *European Planning Studies* 18: 12 (2015) 2073–2081.
A. L. Hughes and L. Palen, "Twitter Adoption and Use in Mass Convergence and Emergency Events," Proceedings of the 6th International ISCRAM Conference (Gothenburg, Sweden, May 2009).
P. Knox and S. Marston, *Human Geography: Places and Regions in Global Context*, 6th edn (Upper Saddle River NJ: Prentice Hall, 2012).
S. Laird, "Super Bowl's First Social Media Command Center an 'Enormous Success,'" *Mashable*, February 7, 2012 <http://mashable.com/2012/02/07/super-bowl-social-media-command-center/> Accessed January 23, 2015.
R. Lee, S. Wakamiya, and K. Sumiya, "Urban Area Characterization Based on Crowd Behavioral Lifelogs over Twitter," *Personal and Ubiquitous Computing* 17 (2013) 605–620.
Y. Long, P. K. Baren, and R. Moore, "The Role of Space Syntax in Spatial Cognition: Evidence from Urban China," Conference Proceedings of the 6th International Space Syntax Symposium, (İstanbul, Turkey, June 2007).
K. Lynch, *The Image of the City* (Cambridge MA: MIT Press, 1960).
K. Lynch, *Good City Form* (Cambridge MA: MIT Press, 1984).
E. Morello and C. Ratti, "A Digital Image of the City: 3D Isovists in Lynch's Urban Analysis," *Environment and Planning B: Planning and Design* 36: 5 (2009) 837–853.
M. Naaman, "Geographic Information from Georeferenced Social Media Data," *SIGSPATIAL Special* 3 (2011) 54–61.
P. L. Pearce and M. Fagence, "The Legacy of Kevin Lynch: Research Implications," *Annals of Tourism Research* 23: 3 (1996) 576–598.
A. Raynsforda, "Civic Art in an Age of Cultural Relativism: The Aesthetic Origins of Kevin Lynch's Image of the City," *Journal of Urban Design* 16: 1 (2011) 43–65.
D. Riff, S. Lacy, and F. Fico, *Analyzing Media Messages: Using Quantitative Content Analysis in Research* (New York: Routledge, 2014).
M. Stellefson, B. Chaney, K. Ochipa, D. Chaney, Z. Haider, B. Hanik, … E. Chavarria, and J. M. Bernhardt, "YouTube as a Source of Chronic Obstructive Pulmonary Disease Patient Education: A Social Media Content Analysis," *Chronic Respiratory Disease* 11: 2 (2014) 61–71.
J. Urry and J. Larsen, *The Tourist Gaze 3.0*, 3rd edn (London: Sage, 2011).
Y.-M. Zhao, A. Zeng, X.-Y. Yan, W.-X. Wang, and Y.-C. Lai, "Unified Underpinning of Human Mobility in the Real World and Cyberspace," *New Journal of Physics* 18: 5 (2016) 053025.
M. Zmudzinska-Nowak, "Searching for Legible City Form: Kevin Lynch's Theory in Contemporary Perspective," *Journal of Urban Technology* 10: 3 (2003) 19–39.

Appendix

Table A1. Spatial references appearing in the Twitter data and used to identify tweets for retention in the sample

Word or phrase	Additional description	IOTC category	Word or phrase	Additional description	IOTC category
116th	Street	Path	Keystone	Street	Path
I69	Interstate highway	Path	Lost & Found	Event service location	Node
I70	Interstate highway	Path	Marriot	Marriot hotel	Landmark
I96	Interstate highway	Path	Massachusetts	Street	Path
Airport		Node	Maxim	Party location	Node
Alabama	Street	Path	Media center	Event venue	Node
Artgarden	Local attraction	Landmark	Meridian	Street	Path
Avenue		Path	Monitoring center	Event venue	Node
Bank One	Downtown building	Landmark	Monument Circle	Downtown landmark	Landmark
Beach bowl	Event venue	Node	NFL Experience	Multi-event venue	Node
Brewhouse	Restaurant & pub	Node	North side		District
Broadripple	Neighborhood	District	Parking lot/garage/ ramp		Node
Canal	Downtown waterway	Edge	Patachou	Downtown cafe	Node
Capitol	Street	Path	Pennsylvania	Street	Path
Carmel	Suburb	District	Playboy	Party venue	Node
Chase	Large downtown building	Landmark	Plaza		Node
Columbia club	Event venue	Node	Pressbox	Event venue	Node
Conrad	Downtown hotel	Landmark	Quad	Downtown areas named on signage	District
Conseco	Event venue	Landmark	Radio row	Event venue	Node
Convention	Convention Center	Landmark	Scotty's	Restaurant & pub	Node
Corner		Node	South side		District
Delaware	Street	Path	St. Elmo's	Downtown restaurant	Node
Downtown		District	Stadium	Event venue	Landmark
East side		District	Station	Downtown breast feeding station	Node
Elbow Room	Restaurant & pub	Node	Street		Path
ExactTarget	Local company office	Node	Stutz	Downtown hotel	Landmark
Fall Creek	River north of downtown	Edge	Super Route	Designated preferred event route	Path
Fieldhouse	Event venue	Landmark	Taproom	Restaurant & pub	Node
Fishers	Suburb	District	Tent	Event venue	Node
Food truck		Node	Tower	Event venue (zipline anchors)	Node
Fountain square		Landmark	Verizon		Landmark
Garage		Node	Village	Super Bowl village	District
Georgia	Street	Path	War memorial	Downtown monument	Landmark
Hilbert	Downtown hotel	Landmark	Washington	Street	Path
Hilton	Local hotel	Landmark	West side		District
Hotel		Node	West Coast Tacos	Fast food vendor	Node
Illinois	Street	Path	White River	River west of downtown	Edge
Intersection		Node	White River State Park	Local attraction	District
Interstate		Path	Zipline	Event attraction	Node