

# Making Sense of Cities Using Social Media: Requirements for Hyper-Local Data Aggregation Tools

**Raz Schwartz**      **Mor Naaman**      **Ziad Matni**  
raz.schwartz@rutgers.edu    mor@rutgers.edu    zmatni@scarletmail.rutgers.edu

Social Media Information Lab  
School of Communication and Information  
Rutgers University  
New Brunswick, NJ 08901

## Abstract

As more people tweet, check-in and share pictures and videos of their daily experiences in the city, new opportunities arise to understand urban activity. When aggregated, these data can uncover invaluable local insights for local stakeholders such as journalists, first responders and city officials. To better understand the needs and requirements for this kind of aggregation tools, we perform an exploratory study that includes interviews with 12 domain experts that utilize local information on a daily basis. Our results shed light on current practices, existing tools and unfulfilled needs of these professionals. We use these findings to discuss the requirements for hyper-local social media data aggregation tools for the study of cities on a large scale. We outline a list of key features that can better serve the discovery of patterns and insights about both real-time activity and historical perspectives of local communities.

## Introduction

Everyday local activities and experiences are now massively shared over various social media platforms through geo-tagged pictures, videos, posts and check-ins. When aggregating these data, an image of a city emerges. This collective representation of the urban habitat can provide local stakeholders better understanding of the city's ongoings. For example, local organizations and professionals such as policy makers, journalists, urban planners and administrators as well as city residents, can use these data to extract invaluable insights about real-time activity (i.e. breaking news) as well as historical perspectives of local communities.

However, it is not yet clear what are the tools needed to make these data useful and accessible. In many cases, users are reappropriating existing tools that were not designed to answer local questions. Since the current efforts and attempts in this area are in their early stages, exploring ways to better format and display these massive volumes of local data to match stakeholders' needs and applications becomes increasingly important.

This work aims to close this gap by creating a detailed list of requirements that is instrumental for developments of hyper-local social media data aggregation tools to the study of cities on a large scale. Inspired by previous projects

(Johansen 1988; Frohlich et al. 2002), we tackle this task through conducting an exploratory study that aims to better understand the requirements for this software by tracing current practices and needs of various users.

To this end, we conducted a set of interviews with a group of journalists, local government employees, administrators and policy makers from various organizations (such as The New York Times, The Daily Beast, and Newark and New Jersey City mayor's offices). Using a series of semi-structured interviews, we asked our participants about their experiences in gathering city-level insights from social media platforms. More specifically, we focused on better understanding our interviewees' current tools and practices as well as their unfulfilled needs and requirements.

In addition to surveying the software and methods used by our participants, we presented them to a set of existing tools and several data visualizations as examples for possible ways to explore and organize the data. We used these examples to query about their usefulness to our participants' current workflows. Based on these interviews, we extract three themes (tracking, analysis and synthesis) that encompass city-level information seeking practices. Following on these themes, we outline a list of data analysis and visualization needs that can inform the design of future systems.

## Related Work

A growing corpora of research in the field of urban computing uses various datasets to better understand the dynamics and patterns of urban activity. For example, several works used mobile phone data to expose travel patterns (Gonzalez, Hidalgo, and Barabasi 2008; Girardin et al. 2008; Isaacman et al. 2010), broad spatio-temporal dynamics (Calabrese et al. 2010; Reades, Calabrese, and Ratti 2009), and social connections (Saramäki and pekka Onnela 2007). Other works used sensors as well as "crowd sensing" information systems to get real-time data about various areas of a city (Roitman et al. 2012). In addition, as more cities decide to open their municipal data sets to the public as part of "open government" initiatives, there is a growing interest in the type of tools and information these data can provide (i.e. City Dashboard (UCL 2012)).

Similarly, with the growing access to geo-tagged data from social media platforms, researches look at ways to gain local insights about city life. Previous projects stud-

ied various aspects of users' activity in cities from tweets, pictures and users check-ins patterns (Crandall et al. 2009; Cheng et al. 2011; Kling and Pozdnoukhov 2012) as well as using these data to identify local events (Becker et al. 2012).

Moreover, considerable work has explored the use of social media data to source for local news and information. These works found that local information sourced from social media tends to be geographically-centered to where that information emanated (Yardi and boyd 2010) and that users communicate informational and helpful messages in times of crisis (Heverin and Zach 2012). Other works found that journalists use social media data to follow-up on breaking news events and source for information shared from the event's scene (Diakopoulos, De Choudhury, and Naaman 2012).

Several research projects explored the use of social media aggregation tools to extract local insights such as the character of various areas of a city (Cranshaw et al. 2012) or prediction of disease transmission (Sadilek, Kautz, and Silenzio 2012). In addition, there are number of commercially available aggregation tools for social media. For example, Geofeedia is geographically-based data aggregation tool (<http://geofeedia.com/>), a web-based service that allows its users to monitor and gather information tagged to a specific geographic area from various social media sources. Other existing tools that provide social media monitoring and alerts in global settings include Dataminr in the financial context (<http://dataminr.com/>) and MarketingCloud for brand management (<http://www.radian6.com/>).

## Methodology

To better understand current practices and requirements for these kind of tools, we conducted a series of interviews with 12 domain experts that utilize local information on daily basis. The semi-structured interviews took place between Dec. 7th, 2012 and Jan. 18th, 2013 and were performed with a group of professionals working at various organizations such as national and local news outlets (The New York Times, The Daily Beast, NewsDay), local government organizations (Newark and Jersey City mayor's offices, the Phoenix Police Department), university administration, and an urban planning institute.

We selected these individuals based on their reliance on local information and the instrumental part this information plays in their daily practice. Moreover, we deliberately did not focus only on one sector (such as journalists and newsrooms) but also solicited feedback from a range of other professions in an effort to represent a spectrum of needs.

The primary goal of our interviews was to explore current practices and tools as well as needs that are missing from existing software. To accomplish this, we developed an interview protocol that focused on our participants daily information gathering methods, past experiences and imagined possibilities. We traced the significance of social media data to their line of work by requesting our participants to describe their daily interactions with social media data. In addition, we asked our interviewees to tell us about particular cases in which these data provided useful local insights. Accordingly, we requested our participants to list and describe

their workflow methods, existing tools as well as shortcomings they encounter.

Moreover, we used several existing aggregation tools and a series of data visualizations examples inspired by current visualization techniques (Figure 1) to enrich the discussion. We utilized these examples as a vehicle to encourage reflection and discussion about the possible applications of these data for their daily practice. Our visualizations examples depicted various ways to display aggregated geo-tagged social media data. In this way, these examples included a heat map visualization plotting data showing the concentrations of information over a city streets map (Figure 1 - top left), a Streamgraph to represent time series data from various social media platforms (figure 1 - bottom left), Cubism graph to display an array of time series data topics (figure 1 - bottom right), and Pack Hierarchy Graph that groups topics based on location and relevance (figure 1 - top right).

All participants were recruited by contacting them directly over email or phone and requesting their professional feedback. The duration of the interviews was approximately 45 minutes on average (no compensation was provided).

Using thematic analysis we examined the interviews for recurring patterns and developed a coding scheme to group our participants' statements. In this way, two researchers examined the transcripts and compared frequencies and co-occurrences of the codes. Based on these results, we identified themes that encompass our interviewees' interaction with city level social media data.

## Results

Based on our results, we identified three human-driven tasks: *Tracking*, *Analysis* and *Synthesis*. Each of these practices portrays a different aspect of handling social media information on a city level. Tracking includes keeping tabs on real time local activity; analysis describes conducting deep investigation; and synthesis refers to extracting meaningful conclusions from aggregated data across various social media platforms. In the following section we will describe each of these practices, explore the ways they are tackled and list unfulfilled needs.

### Tracking

Tracking the city activity in real time was the most common practice mentioned by eleven out of twelve participants. These participants mentioned the ubiquity of social media data and the availability of real-time information as the key aspects that make these data an important resource for local information. Under the theme of Tracking we include detecting events in real time, monitoring ongoing activity as well as instantly reacting to special scenarios.

**Event detection** Our interviewees described various tactics and tools that help them detect special events. For example, five journalists listed sources such as police scanners, news wires and online tools like Google Trends, while two local government employees mentioned local TV channels and news websites. From a social media perspective, eleven of our interviewees considered social media platforms as an

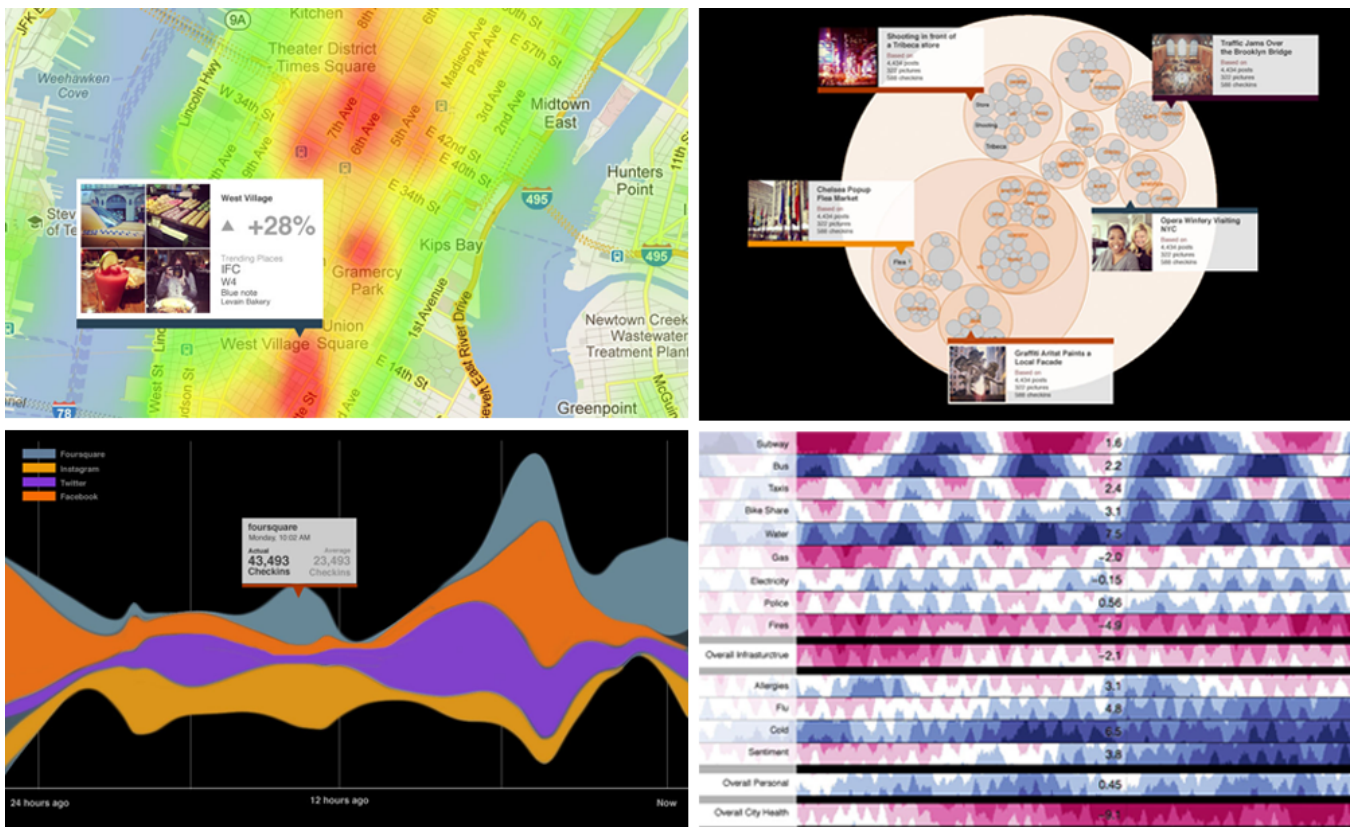


Figure 1: Examples of geo-tagged social media data visualizations mockups. Clockwise: geo-tagged topic groupings, keywords appearance graphs, volume graph and heat map

integral part of their sources to the discovery of local information. As Lisa (media journalist) puts it:

In the old news room you would sit with your police scanner and you would listen to what the police are talking about and know if something's going on. Now I "listen" to social media and I watch the progress of how something is happening and that's very valuable.

Given the profile of the participants we recruited and willingness to participate, it is not surprising that most of them were interested in real time data. The only participant who did not show interest was the urban planner that did not consider real-time data as a significant source to his line of work.

Focusing on data that is unusual was a notion that came up in ten out of our twelve interviews. Our participants emphasized the immense value they could gain from data that displays what is "out of the norm". As Chris (media journalist) explains:

News begins with what's unexpected or what's unusual just monitoring is not useful. I mean, it's useful in the collection of stuff, but seeing what's being monitored right now is not necessarily useful. What's useful is what's different than what's expected, what's unusual.

In this way, finding out first-hand real time information about places that display above than average activity levels

was considered highly valuable. As Joe (media journalist) emphasizes:

Saying that most tweets are coming from the Times Square area is not a surprise to me. I don't really care. But if it was something that was saying this rises above the norm, so that's a thing that I care about. So, if there is something that's happening and there's a frequency that's aberrant...to me that's a hot-spot. It's not about the volume.

As Joe notes, seeing the spread and volume of social media data across the city is not particularly useful. On the other hand, comparing how today's data differs from those of previous days might prove to be significantly useful. Indeed, two interviewees who work in local government organizations noted that detecting abnormality in the data could help them provide better and quicker service to local communities. As Mike (mayor's office employee) explains:

Event detection is super important, so there are certain [events] we would really want to know if people mentioned. And these triggers are super important. So, if someone mentions gunshots, we really want to know that... That's a needle in a haystack that's really important to us... Like, is this thing trending that normally isn't? Or is this hub suddenly overwhelmed? That would be really important to us because we would then



send more police there.

As there are currently no tools that provide this type of event detection, our participants developed various work-around methods that involved keyword searches in real-time using Twitter clients such as Hootsuite and TweetDeck and link services like bit.ly. Five interviewees mentioned using existing tools to perform search based on specific textual strings. For example, Joe (police detective) used names of local gangs to follow tweets about their activity while Lisa (media journalist) chose TweetDeck saved searches to display a real time time line of keywords like “murder” and “shooting”. As she explains:

I always have 6, 7, 8 searches running in my TweetDeck and my interns run searches on the bit.ly site. We are looking for people who are talking about a homicide via one of our search terms then we’ll be able to connect to them, we’ll reach out to them. Sometimes we’re able to get an ID of who the victim is early because people have been talking about it on social media and that comes out in our TweetDeck search.

The use of keywords search provides our participants a way to browse a sub-stream of the information that focuses on a certain issue. As our participants noted, although this method does offer some degree of access to local information, it is limited to a small number of keywords. Moreover, using keywords search retrieves results that are not necessarily related to a certain geographic location. As Chris (media journalist) notes:

What you really want to do is you want to filter it by something that you’re actually paying attention to. To be able to filter, like if you we’re saying, “Ok, show me deviations for the norm in a certain area where shooting is involved”. That would be interesting.

Several interviewees tried to overcome this difficulty by using aggregation tools like Geofeedia that provide search of social media data based on geographical area. However, these participants were not satisfied with the service as it was not updating in real-time and in many cases provided results that were not relevant to their interests.

Finally, when asked about how they would like to find out about the detected events, six participants asked to receive direct alerts (by email, phone etc.) that will draw their attention to the unfolding activity.

**Monitoring events** Seven participants mentioned social media data as source to monitor ongoing activity in the city after an event was detected. In this way, our interviewees follow the live stream of local tweets, pictures, checkins and video clips as they are shared over various social media platforms. Three interviewees mentioned the notion of civic reporting and how people that take pictures on their smartphones can sometimes get a better coverage than that of a reporter. As Chris (media journalists) explains:

I can’t think of anything that broke over social [media] necessarily, but we certainly use it as a good way to monitor when things are breaking. So, when there was the shooting in front of the Empire State Building, for

example, you can monitor that very closely. You can monitor activity around it.

**Interaction with users** In many organizations, monitoring social media on a city level is an active task that requires a designated team of employees to survey the stream of information as well as interact with the users. Three of the local organizations employees we interviewed had specific personnel that were focused on monitoring and engaging with social media users. As Mike (mayor’s office employee) describes:

There is a communication staff and they read all the papers, big news stories about Newark or about the mayor ... For how we manage social media information there are people who monitor various Newark Twitter accounts and there is someone who spends a great deal of their time answering people who tweet at the Mayor, in addition to the Mayor going over the tweets sent to him. He is very responsive.

Our participants noted that interacting with local residents is currently only available through conducting individual search through each of the various social media platforms or as a reply to a user’s contact request or mention. For example, when one of the journalists we interviewed wanted to find sources for a news story, he first searched for tweets that mentioned the topic of his news story and then sent a direct message to users he found their tweets relevant and useful. He regarded this method to be problematic as he felt he was missing a lot of information by people who did not mention his keyword. Similarly, as one of the government employees explained, the people in the mayor’s office are currently only answering messages sent directly to them. They do not actively contact people that complain about city services and therefore: “I imagine a lot of possible interactions with local residents are not materialized”.

## Analysis

In addition to practicing real time tracking of different social media platforms, ten participants described their need to perform an analysis of the data. More specifically, browsing through aggregated flows of information was not sufficient enough. To better understand the information and have the ability to produce local insights such as news and evidence our interviewees had to conduct a deeper investigation through the following tactics:

**Disaggregation** By disaggregating the data our participants can get access to specific tweets, checkins and pictures that are tagged to a certain geographic area. Browsing and having access to content items (such as tweets or pictures) was considered an indispensable part of finding local information. The importance of disaggregating the data was further emphasized when we showed our participants various visualizations that aggregate data in the form of graphs and heat map (Figure 1) and asked them to reflect about the usability of these to their daily work. Brian (media journalist) had a very clear impression about the use of aggregated forms of data:

Often, the most interesting stuff to us is the disaggregated. Easily going from the aggregated to the disaggregated is kinda what you want. Because what you want do is to be able to find out where some interesting stuff is going on, you want to be able to filter out the stuff that is not relevant to that, and then you want to be able to laser-beam focus in on the conversation on a granular level that's relevant and either contact people who are there, contact witnesses, contact people who are saying interesting things, contact experts. That's what you want. So, the aggregated data is interesting, but it's only interesting if you can then get down to the disaggregated data.

Ten out of twelve participants reiterated the importance of having a way to zoom in from the micro, aggregated representation to the micro content item. Having direct access to the source of the information was mentioned as highly valuable both in the process of validating the data as well as following the development of an event.

**Comparison** Having the data archived and then used as a baseline or as a point of reference was another important feature our participants felt does not exist in their current practices. As one of our interviewees noticed, not having access to an historical point of reference felt like “navigating blind-folded” in the social media sphere. And although this kind of comparison is available in various online tools like Google Trends, there is no equivalent to social media data on a city level. As Mike (mayor's office employee):

I think that longer trends are going to be really interesting to us, like, month over month and year over year. Let's just say, it may be almost like Google Flu Trends, where if you can see it over some longer period of time ... if there was something that was indicating something over time that you can see it as a longer short term trend, that might be more useful.

Moreover, Tom, the urban planner we interviewed, considered this kind of city level data to be invaluable for urban planners as it can reveal insights that are missing from sources like census data. Historical data therefore can provide access to local trends as well as residents inputs that would have not been reported otherwise. As Tom explains:

The city planners don't care about live data. Urban policy does not respond quickly. Or nor do we want it to respond quickly. We don't want something that happens right now to result in a new policy. However, adding a searchable databases will be extraordinary useful. Because currently there is no database of this type. But if they could mine that when they need that...

As Tom notes, this kind of data analysis is highly sought after by various local organizations and companies but currently there is no platform that provides this kind of service. Naturally, having access to real time data was less valuable to him as urban planning is not based on short term results but rather studying the city through long term data.

## Synthesis

Using various social media platforms to extract meaningful conclusions in real time was described as a difficult task. After detecting an event and findings the specific information items and sources, our participants have to make the connections between the various tweets, pictures and videos.

**Putting things together** Five of our participants mentioned making connections between various types of information as a highly important part of their work. Our participants accomplished this by combining the results they found on the various social media platforms and then arranging them into a narrative. As Matt (media journalist) explains:

Right when the Aurora shooting was happening we looked for tweets in the area, both tweets that had photos, Instagram, and just text tweets and kind of tried to put together what was happening.<sup>1</sup>

As we can see from Matt's comment, finding the data through a unified stream is not enough. For the data to provide meaningful insights, they must be put together in a way that will help better understand local ongoing. As currently this process is done manually, several participants expressed a need for a cross referencing system that will compare information from various source and integrate them into a single narrative. As Jennifer (media journalist) explains:

Something that would integrate multiple platforms and maps would be very helpful ... anything that would allow you to mine for specifics keyword trends over multiple platforms, that would be great... It makes the task of putting things together maybe a little easier...

**Acknowledging Biases** Five participants emphasized the importance of acknowledging biases when examining and drawing conclusions from the data. Their main concerns were about the biases in the data and the precision of existing methods. More specifically, the demographics of social media users were pointed out as a weak point of the data as it represents only a partial view of the city's residents. In addition, three participants were skeptical about social data mining techniques such as sentiment analysis that were considered imprecise.

## Discussion

Based on our findings, we can now better depict a vision for future hyper-local data aggregation tools. In the following section we will outline a list of requirements and their subsequent features that address current needs and shortcomings in utilizing city level social media data. The three requirements we identified include: *real time event detection*, *identifying and monitoring continuous patterns* and *reviewing and acting on specific content*. These requirements however do not necessarily have to be fulfilled by one specific tool. As various groups of professionals perform different information practices (Leckie, Pettigrew, and Sylvain 1996), a set of applications that cover a range of perspectives may provide better results.

<sup>1</sup>See <http://www.thedailybeast.com/articles/2012/07/20/dark-knight-shooting-tweets-photos-video-from-the-scene.html>

## Real Time Event Detection

Many of our participants were interested in varying types of local events in real time. Hyper-local data aggregation tools can identify these events by tracing deviations from the normal rate of social media data. This type of event detection can play a crucial role in real-time tracking of a city and provide invaluable information for reporters, local government officials and individuals. When detecting these events on a city level, various types should be considered:

1. *Big vs. Small* - Events that include a large group of people vs. events that are limited to a handful of people. For example, a music concert in Central Park and a car accident on the corner of 8 Ave. and 113th st.
2. *Planned vs. Unplanned* - Events that are scheduled and planned beforehand like a presidential visit to the 9/11 memorial vs. spontaneous, unorganized activity such as Occupy Wall Street demonstrations.
3. *Ongoing vs. Finished* - Following events in real time as they unfold vs. reporting and monitoring the results of an event.

In addition to these characteristics, events can also be associated with different categories such as traffic, weather and criminal activity. As various users would want to pay attention to different events, some tools could be specific for a certain type of event or category, while others could learn the users's preferences, or allow the users to manually set their preferences. For example, local police officials could find out about broken traffic lights if they set their preferences to "small events" and the category to "traffic". As a result, real time detection may help these professionals to swiftly act and provide assistance.

**Alerts** Tracking live data requires vast attention. To overcome this difficulty, hyper-local data aggregation tools could include an alerts system. A customized live alerts system can draw the users' attention to any abnormal activity that they set as relevant for them. Users should be able to set the frequency of these notifications as well as their format (i.e. text messages, emails or even a "a flashing light", as one of our participant suggested).

## Identifying and Monitoring Continuous Patterns

As our interviewees noted, studying the city through social media data offers a way to better understand both people's activity and interests. When researching city level data, we can identify continuous patterns of both physical and online activity. For example, we can examine mobility patterns by looking at users check-ins activity over several months or trace local sentiment by studying the contents of geo-tagged tweets. The following generic requirements can facilitate the identification and investigation of these continuous patterns.

**Real-time data vs. Historical data** Archiving historical social media data of a city can provide a crucial baseline for the discovery of patterns. The ability to compare current volumes and spread of social media data across the city with past statistics can offer a more detailed account to explain deviations in the data. Studying historical social media

data may also reveal hidden patterns of local communities such as transportation routes and areas usage by time of day. Access to historical perspectives of city level data can provide urban planners and policy makers significant insights and evidence needed in the processes of resources allocations among local communities. For example, examining the hourly change in the number of people that frequent public spaces such as Times Square can surface valuable insights for law enforcement forces and sanitation services professionals.

**Real-Time Aggregated Representation** Hyper-local data aggregation tools can help monitor patterns of people's activity in the city by providing a real-time aggregated representation of the data. Creating a live representation of the city from real-time data requires special attention to the spatio-temporal characteristics of the data. For example, plotting the data over a city map that automatically zooms in and out of areas that show continuous traffic patterns can provide a way to orient the user and perform sensemaking practices. Similarly, using real-time aggregated representation, users can monitor ongoing temporal patterns such as check-ins in bus stops during rush hour.

## Reviewing and Acting on Specific Content

In many cases, our participants wanted a way to review specific pieces of information. By examining the data and gathering various information items, users can act and produce relevant conclusions. These conclusions could be part of a police investigation for evidence, a journalist inquiry into a local story, or a mayor's office survey of citizens complaints.

**Drilling Down** To obtain a deep understanding of the data, it is crucial for users to have the ability to "drill down" and view specific information items. In this way, moving from a bird's eye view to a specific content item can help users to validate the data as well as find a detailed account to answer their questions. Similarly to the visual-information-seeking mantra: "Overview first, zoom and filter, than details on demand" (Shneiderman 1996), this action can be accomplished by focusing on topics and/or geographical areas that are relevant for the users.

*Filter by topic:* Although users currently find their information through keywords search, this does not have to be the case. Hyper-local data aggregation tools could group and present information items based on relevancy to a specific topic. For example, if a journalist covers a specific topic like traffic, these tools could provide a full account of the relevant information that fits into this topic (such as car accidents, traffic jam, bike share). Building on historical data, these tools can learn automatically how to associate different content items into topic grouping like weather, crime and sports. One of the challenges here is to identify what are the topics that actually relate to local activity. In other words, if users tweet about shooting, the system could identify that these reportings relates to a local incident not one that took place in a different city hundreds of miles away.

*Zoom to areas:* As geo-tagged social media content items carry exact longitude and latitude details, plotting them over a map can show activity up to a street corner level. These



kind of mappings already exist in current tools but they are limited by a geographical bounding box. This limitation creates a critical gap in access to data as users can not review information from outside the box as well as the amount of retrieved items is limited. Hyper-local data aggregation tools therefore could offer a dynamic depiction of an area data that is not limited to certain borders and not constrained to a simple depiction of dots on a map. In other words, users should be able to move freely across a geographical region and review the data in meaningful ways such as by visualizing the networks of relations between the various content items. For example, all items that were shared by the same user or their followers will be marked differently thus unveiling local communal activity.

**Comprehensive Review** These tools could provide access to the wide range of content types (text, pictures and videos). As various social media platforms allow their users to produce and share a range of these different types of information, the city data becomes more diverse and meaningful. Consequently, users are displayed with a more wholesome view of local activity.

In addition to the various types of content, hyper-local data aggregation tools could also provide a unified access to the different types of platforms. Combining information from various social media networks and presenting it in one place is critical for data aggregation tools. In other words, when creating a combined presentation of the data, these tools could offer an arena to consolidate and review all social media data along with open local government data that is relevant to the study of the city. For example, police detectives can use combined information from Twitter, Instagram, Foursquare and the city's crime incidents data to cross reference information items and extract a more detailed account of a crime scene.

In many cases, when following various social media platforms, it is not clear which platform can provide better results. Hyper-local data aggregation tools can therefore point users to the platform that will be the most relevant to the type of information the users are seeking. In other words, if we are looking for information about a fire in SoHo, the system will prioritize sources that show a stronger signal of relevant data in that context.

**Reaching Out** In addition to disaggregating the data and accessing an individual tweet, picture or the user details, these tools could also include a built in option to contact and interact with the individual user. These tools may include practices such as responding, posting, commenting and sending a direct message to a specific user. Utilizing this communication channel, users like reporters and local government employees can ask for further information and engage in conversation with the city's residents.

## Conclusion

In this work we studied how city stakeholders such as local newsrooms, mayor's offices, police departments as well as urban planners are currently making sense of social media information. Based on their feedback, we traced what are

the requirements that need to inform future development of tools. As we show, developing tools that focus on detecting abnormal activities, monitoring ongoing events, performing meaningful review and then providing ways to act accordingly, can provide invaluable information that can improve the process of making sense of our cities.

The prevalence of publicly available and location-annotated social media data promises to shift the manner in which we think about, and perceive, the urban habitat. As more people document their daily activities and interactions in the city, hyper-local data aggregation tools would be able to extract urban insights that were not easily discoverable before. Utilizing social media tools to the study of cities can therefore prove to be invaluable in our goal to better understand the dynamics of our cities.

## References

- Becker, H.; Iyer, D.; Naaman, M.; and Gravano, L. 2012. Identifying content for planned events across social media sites. In *Proceedings of the 5th ACM international conference on Web search and data mining*, WSDM '12, 533–542. New York, NY, USA: ACM.
- Calabrese, F.; Pereira, F. C.; Di Lorenzo, G.; Liu, L.; and Ratti, C. 2010. The geography of taste: analyzing cell-phone mobility and social events. In *Proceedings of the 8th international conference on Pervasive Computing*, Pervasive '10, 22–37. Berlin, Heidelberg: Springer-Verlag.
- Cheng, Z.; Caverlee, J.; Lee, K.; and Sui, D. Z. 2011. Exploring Millions of Footprints in Location Sharing Services. In *Proceedings of the 4th International Conference on Weblogs and Social Media*, ICWSM '11. Menlo Park, CA, USA: AAAI.
- Crandall, D. J.; Backstrom, L.; Huttenlocher, D.; and Kleinberg, J. 2009. Mapping the world's photos. In *Proceedings of the 18th international conference on World wide web*, WWW '09, 761–770. New York, NY, USA: ACM.
- Cranshaw, J.; Schwartz, R.; Hong, J. I.; and Sadeh, N. M. 2012. The livelihoods project: Utilizing social media to understand the dynamics of a city. In *Proceedings of the 6th International Conference on Weblogs and Social Media*, ICWSM '12. Menlo Park, CA, USA: AAAI.
- Diakopoulos, N.; De Choudhury, M.; and Naaman, M. 2012. Finding and assessing social media information sources in the context of journalism. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, CHI '12, 2451–2460. New York, NY, USA: ACM.
- Frohlich, D.; Kuchinsky, A.; Pering, C.; Don, A.; and Ariss, S. 2002. Requirements for photoware. In *Proceedings of the 2002 ACM conference on Computer supported cooperative work*, CSCW '02, 166–175. New York, NY, USA: ACM.
- Girardin, F.; Fiore, F. D.; Ratti, C.; and Blat, J. 2008. Leveraging explicitly disclosed location information to understand tourist dynamics: a case study. *Journal of Location Based Services* 2(1):41–56.
- Gonzalez, M. C.; Hidalgo, C. A.; and Barabasi, A.-L. 2008. Understanding individual human mobility patterns. *Nature* 453(7196):779–782.

- Heverin, T., and Zach, L. 2012. Use of microblogging for collective sense-making during violent crises: a study of three campus shootings. *Journal of the American Society for Information Science and Technology* 63(1):34–47.
- Isaacman, S.; Becker, R.; Cáceres, R.; Kobourov, S.; Rowland, J.; and Varshavsky, A. 2010. A tale of two cities. In *Proceedings of the 11th Workshop on Mobile Computing Systems and Applications*, HotMobile '10, 19–24. New York, NY, USA: ACM.
- Johansen, R. 1988. *GroupWare: Computer Support for Business Teams*. New York, NY, USA: The Free Press.
- Kling, F., and Pozdnoukhov, A. 2012. When a city tells a story: urban topic analysis. In *Proceedings of the 20th International Conference on Advances in Geographic Information Systems*, SIGSPATIAL '12, 482–485. New York, NY, USA: ACM.
- Leckie, G. J.; Pettigrew, K. E.; and Sylvain, C. 1996. Modeling the information seeking of professionals: A general model derived from research on engineers, health care professionals, and lawyers. *The Library Quarterly* 66(2):161–193.
- Reades, J.; Calabrese, F.; and Ratti, C. 2009. Eigenplaces: analysing cities using the space - time structure of the mobile phone network. *Environment and Planning B: Planning and Design* 36(5):824–836.
- Roitman, H.; Mamou, J.; Mehta, S.; Satt, A.; and Subramaniam, L. 2012. Harnessing the crowds for smart city sensing. In *Proceedings of the 1st international workshop on Multimodal crowd sensing*, CrowdSens '12, 17–18. New York, NY, USA: ACM.
- Sadilek, A.; Kautz, H.; and Silenzio, V. 2012. Predicting disease transmission from geo-tagged micro-blog data. In *26th AAAI Conference on Artificial Intelligence*, volume AAAI '12. Menlo Park, CA, USA: AAAI.
- Saramäki, J., and pekka Onnela, J. 2007. Structure and tie strengths in mobile communication networks. *Proceedings of the National Academy of Sciences of the USA* 104(18):7332–7336.
- Shneiderman, B. 1996. The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings of the 1996 IEEE Symposium on Visual Languages*, VL '96. Washington, DC, USA: IEEE Computer Society.
- UCL. 2012. City dashboard, casa research lab. <http://citydashboard.org/london/>.
- Yardi, S., and boyd, d. 2010. Tweeting from the town square: Measuring geographic local networks. In *Proceedings of the 4th International Conference on Weblogs and Social Media*, ICWSM '10. Menlo Park, CA, USA: AAAI.