

Tracking Community Development from Social Media

Mohammed Ghriga

*Department of Technology, Innovation and Computer Science
Long Island University Brooklyn
Brooklyn, New York, USA*

Mohammed.Ghriga@liu.edu

Richard Shang

*Department of Technology, Innovation and Computer Science
Long Island University Brooklyn
Brooklyn, New York, USA*

Di.Shang@liu.edu

Rshana Shurriah

*Department of Technology, Innovation and Computer Science
Long Island University Brooklyn
Brooklyn, New York, USA*

Rshana.Shurriah@liu.edu

Abstract

The use of social media and networking has become a significant community development tool to reach out to a wide audience, share information in real-time, improve communications, and build relationships. In this research, we aim to demonstrate how social media analytics provide a valuable set of techniques for deriving community development information from social media. We examine and provide an inside look at the activities and structure of a local community - Brooklyn Tech Triangle based on the data collected from Twitter on the key advocates of Brooklyn Tech Triangle. Our preliminary analysis shows that Tech Triangle community development practitioners use social media to promote the fundamental tenets of the community and communicate their initiatives. In addition, they keep leadership roles in connecting the community and some value-added influencers have very important influence on the transfer of information through the social network.

Keywords: social media, social network, online community, information systems development, data analytics

1. Introduction

The use of social media and networking has become a significant community development tool to reach out to a wide audience, share information in real-time, improve communications, and build relationships. Social media data provide the opportunity for understanding how a community is formed and maintained online. Community development practitioners can actively engage in learning and understanding about community issues with information on social media. Social media and networking can enhance the leadership capacity of community members, leaders, and groups within the community through networking. Information can be more broadly disseminated and greater transparency and accountability can lead to increases in trust within community members [3][6]. In this research we aim to demonstrate how social media analytics provide a valuable set of techniques for deriving community development information from social media. We examine and provide an inside look at the activities and structure of a local community - Brooklyn Tech Triangle.

Brooklyn is the largest of New York City's five boroughs by population and the second largest by area. The Brooklyn Tech Triangle made up of the following key partners: Downtown

Brooklyn, DUMBO, and the Brooklyn Navy Yard, as shown in Figure 1. It has become a magnet for the world's pioneering, energetic, and creative entrepreneurs and has emerged as New York City's largest cluster of tech activity outside of Manhattan. Much of this growth can be attributed to the fact that innovation firms want to be close to each other to share ideas and benefit from a diversity of expertise. The Brooklyn Tech Triangle initiative is an effort to bring together support for growing the innovation economy of the borough. Since the use of social media and social networking has become a significant force in social interaction and economic development, the Brooklyn Tech Triangle along with its key partners all have their individual social media presences to promote an active and cohesive community, attract entrepreneurs and investment, and connect the local community with the economic opportunity and resources.

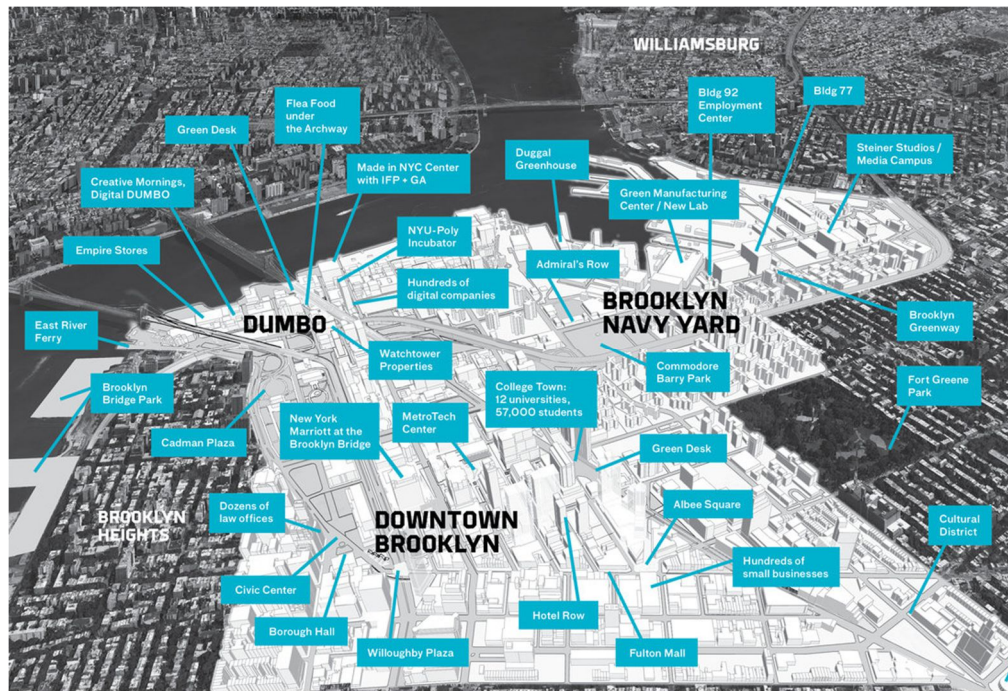


Fig 1. Map of Brooklyn Tech Triangle and its key resources (Image by Brooklyn Tech Triangle [4])

Based on the data collected from Twitter over the span of 21 months on the key players of Brooklyn Tech Triangle, we aim to investigate specifically two aspects of the community development:

1. How the community development practitioners use social media to promote the fundamental tenets of the community and influence how participants act?
2. Whether and how the social network enhance the leadership capacity of community development practitioners and other members through networking?

Our preliminary analysis shows that Tech Triangle community development practitioners promote an active community centered in their initiatives and keep leadership roles in connecting the local community with the economic opportunity and resources. In addition, value-added influencers play important roles in the community and have an influence on the transfer of information through the social network.

2. Theoretical Background

For the last few years, interest in social networks has increased and become an active research area mainly due to commercial and social importance, the potential of applications, and the value of analysis. As a result, new algorithms, models, and applications have been developed

to address particular domains and resolve distinct problems in social media [8]. Data analysis and its measurements are considered to be integral components of network research [11].

Organizations understand the importance of utilizing social media to exist in today's technologically advanced market. From social media, companies can have first-hand knowledge about their potential customer's opinions, brand image, and what they are saying about their brand. Organizations nowadays focus on analyzing social networks to know who is central in these networks and identifying the members of the social network community who are best placed to diffuse information. For instance, analyzing twitter data helps institutions or companies to have better understanding about their customers' feedbacks and opinions, attitudes, perceptions and behavior [10, 11]. To conduct such information by the help of traditional survey or interview methods could be a costly troublesome process.

Social media and networking provide a free and ubiquitous method of communicating, sharing of information, and involvement of a diverse constituency in community development efforts [16]. Mining social network media such as twitter for attitudinal data can be a better process for generating insights and knowledge [13]. In addition, organizations and policy makers can benefit from targeting members to effectively spread valuable information [2][5].

3. Datasets and Analysis

Table 36. Twitter data of Brooklyn Tech Triangle.

Key Partner	Search Term	Number of Tweets
Tech Triangle	TechTriangleU, #BKTechTriangle	262
Downtown Brooklyn	DowntownBklyn, #DowntownBrooklyn	1559
Dumbo	DumboBid	608
Navy Yard	BklynNavyYard, BLDG92, #BrooklynNavyYard	688

The Brooklyn Tech Triangle along with its key partners: Downtown Brooklyn, DUMBO, and the Brooklyn Navy Yard all have their individual websites and social media presence to promote community development and organize events. Using related Twitter accounts and hashtags (as shown in Table 1), we searched and collected data through IBM dashDB on IBM's cloud platform as a service – IBM Bluemix. Total 3117 tweets were collected from Twitter over the span of 21 months, from April 2014 to December 2015. Data were then analyzed and visualized on BlueMix with *Insights for Twitter* and R packages: *RBGL*, *SNA*, *TM*, and *Wordcloud*.

There are many indicators of successful online communities. At this stage, we use the following metrics to track the development of Brooklyn Tech Triangle community.

1. The engagement of community development practitioners and users.
2. Time to respond and dissemination of community information.
3. The sentiment of community users over time.
4. Influence of community development practitioners and active users.

4. Results

We first examine the engagement of community development practitioners in using Twitter to promote the fundamental tenets of the community and influence the topics that users most talked or followed. As shown in Figure 2, the community development practitioners are the most active in term of content and traffic generation in the community. The number of tweets posted or shared by the practitioners accounts for more than 20% on average of the total number of tweets generated by the community over time. Moreover, large portion (more than 50%) of these tweets are original posts, which indicates active engagements of contributors.

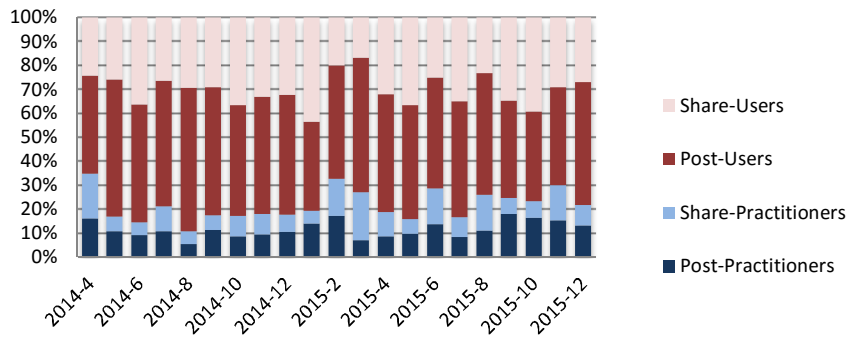


Fig 2. Tweets contributed by of practitioners and users in percentage

Figure 3 illustrates the word cloud analysis of our dataset. Word clouds are graphical representations of word frequency that give greater prominence to words that appear more frequently in a source text. The larger the word in the visual the more common the word was in the documents.

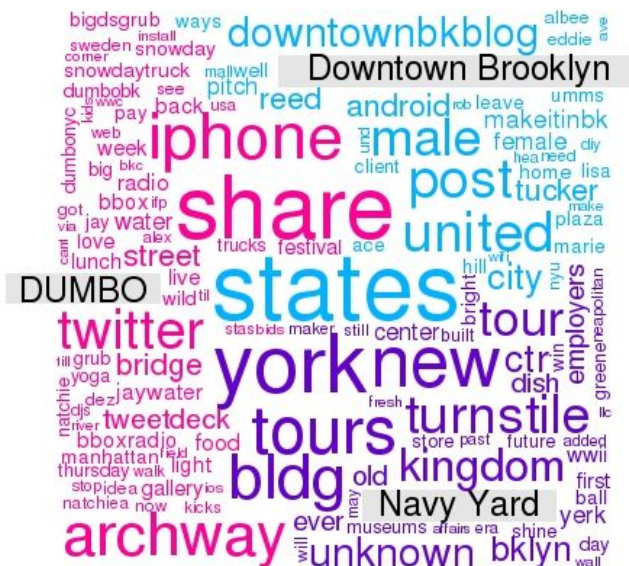


Fig 3. Comparison cloud of tweets about Brooklyn Tech Triangle

The *Downtown Brooklyn* undertakes diverse activities include attracting new investment and improving the environment for existing companies; facilitating infrastructure development that promotes an active and cohesive community. As shown in Figure 3, the topics that was most talked about for Downtown Brooklyn during the period included 'Makeitinbk', 'Brooklynchamber', 'nyupoly; 'tucker', 'metrotech'. For Example, 'Makeitinbk'- *Make It in Brooklyn*, was one of the large words in the visual in Downtown Brooklyn's word cloud, and it fits in with Downtown Brooklyn's initiative to connect with Brooklyn's best and brightest by hosting Innovation Summit and build entrepreneurs Brooklyn network by hosting pitch contests [14]. Figure 4 shows the number of retweets within 24 hours of the announcement of a *Makeitinbk* pitch contest on October 5, 2015. The announcement attracted more than 3000 retweets within 24 hours, and it suggests that community users quickly respond to community events and disseminate the information over the community. However, the number of retweets sharply decreased after 24 hours, which also reflects the fact that trending topics on Twitter can quickly change from one term to another [12].

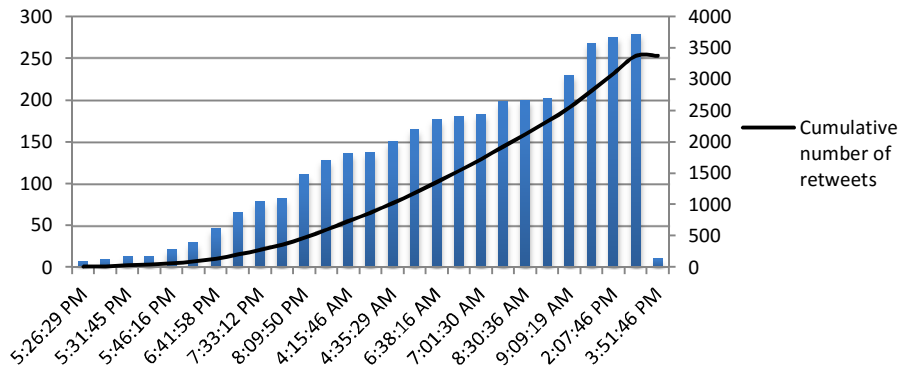


Fig 4. Number of retweets within 24 hours of the announcement of events

The *DUMBO Improvement District* is a non-profit organization founded to enhance and promote the quality of life in DUMBO, Brooklyn. The area has emerged as one of New York City's premier arts districts, with a cluster of art galleries. DUMBO, a neighborhood located between the Brooklyn and Manhattan Bridges, is Brooklyn's epicenter of art and creativity [7]. The DUMBO Twitter handle shares information, news and events on behalf of the businesses in the area. Therefore, its word cloud presents more art gallery and food vendor related topics. In Figure 3, large words in the word cloud, 'dumbolot', 'artdumbobk', 'newyorkhabitat', 'artindumbo', 'nymediacenter', speak directly to DUMBO's initiatives. Art in Dumbo, was one of the large words in the visual for DUMBO and fits in with its mantra of being epicenter of art and creativity.

The *Brooklyn Navy Yard Development Corporation* is a not-for-profit corporation that manages the Brooklyn Navy Yard on behalf of the City of New York. Its mission is to fuel New York City's economic vitality by creating and preserving quality jobs, anchoring the City's modern industrial sector and businesses, and connecting the local community with the economic opportunity and the resources of the Yard. Through investments in new green buildings and infrastructure, the Yard has become a national model for sustainable urban industrial parks and home to modern manufacturers that are tech-driven and socially responsible. In Figure 3, the topics that was most talked about for the Brooklyn Navy Yard during the period included 'employers', 'workers', 'exhibit', 'work', 'intern', which all fit into the Navy yards initiatives.

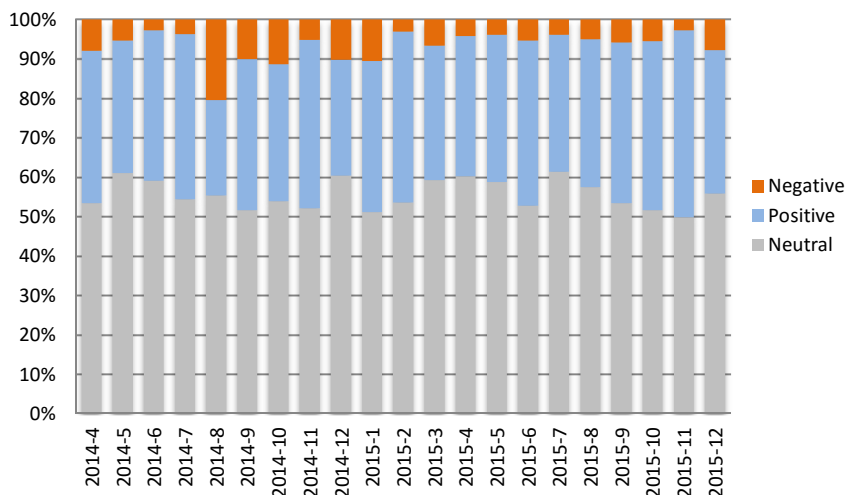


Fig 5. Sentiment of tweets over time

Social media sentiment analysis can be an excellent source of information and can provide insights that can improve campaign success and customer services. Using opinion lexicon by [9], we conducted sentiment analysis of the Twitter data, and our results (as shown in Figure 5) suggest that community users generally express positive sentiment during the period of time.

We next examine whether the community development practitioners present leaderships in the community and enhance the leadership capacity through networking. Knowing who is influential or central in a community is important for the community members as well as for the businesses and policy makers [2].

The most influential users in the network are labeled in the pruned network graph (Figure 6) based on Brandes' betweenness scores. Finding out centrality score or central node is important because it could help spread information in the social network faster, and it could also help protect the network from breaking [1]. A node with high betweenness centrality has a large influence on the transfer of items through the network, under the assumption that item transfer follows the shortest paths [15]. As shown in Table 2, all of *Downtown Brooklyn*, *DUMBO* and *Navy Yard* ranked top in the list reflecting their leadership and influence on the social network.



Fig 6. Network analysis of Twitter data of Brooklyn Tech Triangle

In addition to the practitioners, certain Twitter accounts in the network play important roles in keeping the community active and the transfer of information through the network. For example, *TuckerDBP*, an advocator, is active in tweeting and blogging about the development of Brooklyn; and *SFC*, a college's entrepreneurship center, promotes entrepreneur and practitioner engagement with resources through its Open Virtual Incubator. Those value-added influencers promote the community development initiative and enhance the leadership of the community development practitioners.

Table 2. Centrality scores of key accounts

Account	Rank	Score
DowntownBklyn	1	190966.7
DUMBOBID	2	118769.4
TechTriangleU	3	40395.7
BLDG92	4	29458.0
Brownstoner	6	14137.4
Explore Brooklyn	7	11372.7
SFC Entrepreneur	8	11864.3
makeitink	9	10766.3

5. Discussion

Social media analytics can generate important insights to researchers and practitioners to track community development and recommend communication strategies. Our analysis shows that the Tech Triangle community development practitioners use social media to promote the fundamental tenets of the community and communicate their initiatives. In addition, they keep leadership roles in connecting the community and those value-added influencers have an influence on the transfer of information through the social network. Those findings suggest that community development practitioners should identify active influencers in the network and utilize them to enhance their leadership and quality information sharing and communication.

Techniques used in our research and some of our findings can be useful for practitioners to design platforms that use social media as a means of community development tracking. Such a system could perform tracking task with minimal supervision and with the ability to provide time-critical alerts and decision-support to community development practitioners. Social media data can be collated from different data sources and stored in a data repository. Data relevant to a particular interest can be retrieved, and different analysis techniques (namely, topic analysis, sentiment analysis, and influence analysis) would operate on the retrieved data. The output of these analyses will be combined together in a trend analysis to track emerging or prevalent trends over time. An insight generation and reporting system will consolidate different forms of analytic results together to produce reports pertaining to insights for identifying key influencers, tracking early signs of new topics, or assessing community users' sentiments and concerns over issues.

6. Conclusion

Using the case of Brooklyn Tech Triangle, we demonstrate how social media analytics provides a valuable set of techniques for deriving community development information from social media. While social media is considered a less formal platform for rapid release of critical information, our analysis provide evidence that community development practitioners can potentially benefit from leverage social media as new tools to improve their communication strategies. In addition, our approach and the techniques used can apply to the analysis of other communities and help practitioners to design platforms using social media to track community development.

Our findings are based on analysis of a local community that has a relative small Twitter data set. This could cause our findings not to be generalized to other communities. We can overcome the limitation in our future work by investigating more communities and also include data from other social media sources. It is worth mentioning that tracking community

development using social media also presents many challenges, such as constant monitoring and real-time analysis of large influx of data, high level of background noise. However, the underlying value coupled with challenges of using social media warrants future research.

References

1. Bakshy, E., Hofman, J.M., Mason, W.A, Watts, D.J.: Everyone's an Influencer: Quantifying Influence on Twitter. In: Proceedings of the 4th ACM International Conference on Web Search and Data Mining, pp. 65–74. ACM, New York, NY, USA (2011)
2. Banerjee, A., Chandrasekhar, A.G., Duflo, E., Jackson, M.O.: Gossip: Identifying Central Individuals in a Social Network. MIT Department of Economics Working Paper No. 14-15 (2014)
3. Benevenuto, F., Rodrigues, T., Cha, M., Almeida, V.: Characterizing User Behavior in Online Social Networks. IMC'09, Chicago, Illinois, USA (2009)
4. Brownstoner, <http://www.brownstoner.com/real-estate-market/brooklyn-tech-triangle-needs-3-7-million-more-square-feet-of-offices>. Accessed July 1, 2016
5. Cha, M., Haddadi, H., Benevenuto, F., Gummadi, K.: Measuring User Influence in Twitter: The Million Follower Fallacy. In: Proceedings of the 4th International AAAI Conference on Weblogs and Social Media (ICWSM), George Washington University, Washington DC (2010)
6. Chau, M., Xu, J.: Business Intelligence in Blogs: Understanding Consumer Interactions and Communities. MIS Quarterly, 36(4), pp. 1189-1216 (2012)
7. Graham, J.: It's Hip to be Tech in Brooklyn's Dumbo. May 7, USA TODAY (2013)
8. Hamed, A.: An Exploratory Analysis of Twitter Keyword Networks and Knowledge Discovery Applications. Graduate College Dissertations and Theses, Paper 325, The University of Vermont (2014)
9. Hu, M., Liu, B.: Mining opinion features in customer reviews. In: Proceedings of the 19th national conference on Artificial Intelligence, pp. 755-760 (2004)
10. Jansen, B. J., Zhang, M., Sobel, K., Chowdury, A.: Twitter Power: Tweets as Electronic Word of Mouth. Journal of the American Society for Information Science and Technology. 60(11), pp. 2169-2188 (2009)
11. Kolaczyk, E.D., Csárdi, G.: Statistical Analysis of Network Data with R. Springer Science + Business Media New York, Springer-Verlag, New York (2014)
12. Lin, J., Mishne, G.: A Study of “Churn” in Tweets and Real-Time Search Queries. In: Proceedings of the 6th International AAAI Conference on Weblogs and Social Media (2012)
13. Luong, T.T., Houston, D.: Public Opinions of Light Rail Service in Los Angeles, an Analysis Using Twitter Data. iConference 2015 paper, iSchool, University of Illinois at Urbana–Champaign (2015)
14. Makeitinbk, <http://www.makeitinbk.com>. Accessed July 1, 2016
15. Poulin, R., Boily, M.C., Masse, B.R.: Dynamical Systems to Define Centrality in Social Networks. Social Networks, 22, pp. 187-220 (2000)
16. Scanfeld, D., Larson, E.L.: Dissemination of Health Information through Social Networks: Twitter and Antibiotics. American Journal of Infection Control, 38(3), pp. 182-188 (2010)