Understanding Online Civic Engagement: A Multi–Neighborhood Study of SeeClickFix

Christopher Yong[†], Charalampos Chelmis[†], Wonhyung Lee[⋆] and Daphney–Stavroula Zois[‡]
[†]*Computer Science Department*, [⋆]*School of Social Welfare*, [‡]*Department of Electrical and Computer Engineering*,
University at Albany, SUNY, Albany, NY, USA
Emails: {cyong, cchelmis, whlee, dzois}@albany.edu

***Abstract*—The relationship between local governments and the general public is being redefined by the increasing use of online platforms that enable participatory reporting of non–emergency urban issues, such as potholes and illegal graffiti by concerned citizens to their local authorities. In this work, we study, for the first time, participatory reporting data together with neighborhood–level demographics, socioeconomic indicators, and pedestrian friendliness and transit and bike scores, across multiple neighborhoods in the Capital District of the New York State. Our data–driven approach offers a large–scale, low–cost alternative to traditional survey methods, and provides insights on citizen participation and satisfaction, and public value creation on such platforms. Our findings can be used to guide government service departments to work more closely with each neighborhood to improve the offline and online communication channels through which citizens can report urban issues.**

***Index Terms*—web and society, urban science, smart cities, e-society, mixed methods, computational social science**

I. Introduction

Civic engagement platforms, such as SeeClickFix [1] and FixMyStreet [2], and "Government 2.0" applications [3], [4] have revolutionized the way citizens interact with their local governments. Under this model, citizens can report and resolve non–emergency urban issues in addition to collecting, analyzing and sharing knowledge about their local environment (e.g., air quality [5], fuel consumption [6], and bus arrival times [7]).

The possibility to be heard, and the ability to actively shape the urban spaces they reside, provides citizens with a strong intrinsic motivation to participate in urban issue reporting [1], [8], [9]. Administrative bodies and policy makers can utilize such reporting to create public value by improving the response speed of public services and enhancing citizens–government trust [10]. According to the social exchange theory however [11], local governments must meet citizens' needs to ensure their continuous participation and engagement. Nevertheless, the frequency of use of participatory technologies and the perceived benefit of using them from citizens may depend on the characteristics of citizens and the neighborhood environments with which they are engaging [12]–[14].

Herein, we shed light on the characteristics of online participatory urban issue reporting as a function of the real (i.e., off–line) world. Our work contributes to the online social network analysis and mining community in the following ways:

- We collect and analyze for the first time data from the SeeClickFix[1] civic engagement platform in conjunction with neighborhood–level socioeconomic and demographics indicators from Statistical Atlas[2], and pedestrian friendliness scores from WalkScore[3].
- We make our data findable, accessible, interoperable, and reusable[4] by (i) making it retrievable through an open and free, globally unique and eternally persistent identifier at https://doi.org/10.7910/DVN/WQ2M1H provided by Dataverse[5], (ii) including rich citation terms[6] and social sciences metadata[7], compliant with Dublin Core[8], and (iii) opening it to the public domain using a CC0 – "Public Domain Dedication" license.
- We list open research questions to catalyze future work.

II. Related Work

While community–focused social media and civic engagement platforms have appeared relatively recently, various studies have been performed to quantify social capital and measure citizen participation [15]–[17]. Prior work, however, focuses mainly on political and social activities [17] with research in the context of civic engagement platforms being in its infancy.

III. Methodology

A. Site Selection

We focus on Albany County, New York, the central core of the Capital District of New York State, which ranks among the

This material is based upon work supported by the National Science Foundation under Grant No. ECCS–1737443.

ASONAM '19, August 27–30, 2019, Vancouver, BC, Canada

ACM ISBN 978-1-4503-6868-1/19/08...$15.00
http://dx.doi.org/10.1145/3341161.3345330

[1]http://en.seeclickfix.com
[2]https://statisticalatlas.com/United-States/Overview
[3]https://www.walkscore.com/
[4]https://www.force11.org/group/fairgroup/fairprinciples
[5]https://dataverse.org/
[6]http://dublincore.org/documents/dcmi-terms/
[7]https://github.com/IQSS/dataverse/blob/master/scripts/api/data/metadatablocks/social_science.tsv
[8]http://dublincore.org/

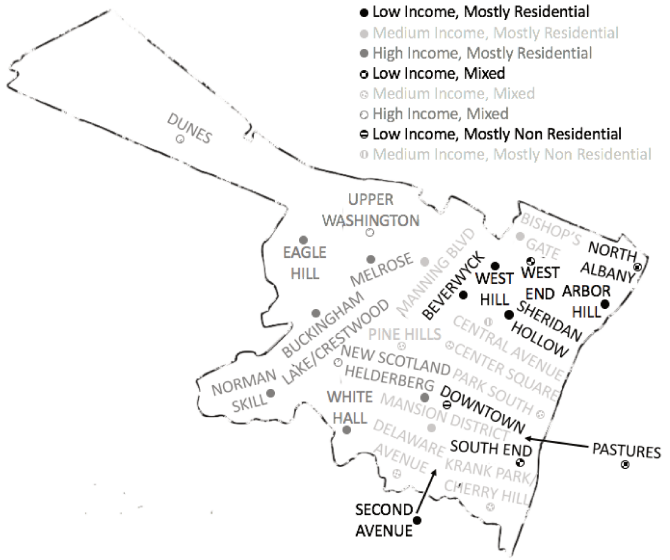


Fig. 1: Neighborhoods in Albany, color–coded by income as Low– ($\$17,973 - \$34,671.48$), Medium– ($\$34,671.49 - \$53,423.87$), or High–Income ($\$53,423.88 - \$77,804.62$) and type (Mostly Residential, Mixed, or Mostly Non Residential).

top "*most engaged and most responsive local governments*" withing cities in which SeeClickFix is actively used and whose population rages between $50-150$k[9]. Additionally, urban issue reports for Albany are more descriptive (~ 1.8 words per reported issue) than for most major cities (e.g., ~ 0.03 words per reported issue in Chicago). Being the 4th largest metropolitan region in the New York State and the 45th largest in the United States, Albany is representative of any typical city in the United States. Albany is a multiracial and multiethnic city that contains nearly $99,000$ residents (on average 52% White, 29% Black or African American, 9% Hispanic or Latino, 6% Asian, and growing refugee populations)[10], and contains a mixture of income–level and type neighborhoods (see Figure 1).

B. Reported Urban Issues from SeeClickFix

The SeeClickFix civic engagement platform bridges the communication gap between residents and their local governments by allowing citizens to simply take "*a photo of a pothole or other problem, geo–locate it and hit submit. SeeClickFix publicly documents the issue and notifies local governments and others who resolve the problem.*" [1]. When an issue is first reported on SeeClickFix, it is labeled as "Open", waiting to be handled by a city official. To signify that the reported issue has been diverted to the appropriate agency for resolution (or that the issue has been resolved), a verified account associated with a city official comments back and marks the issue as "Acknowledged" (similarly "Closed"). In most cases, reported issues remain "Closed", however, a "Closed" issue can be marked as "Open" by the reporter (or another user) to inform the local government about an unresolved (or potentially recurrent) urban issue. The online authority may again mark the issue as "Acknowledged" (similarly "Closed"), and the cycle may theoretically repeat indefinitely (see Figure 2). Reported issues are also assigned to categories (out of 35 categories in total), broadly divided into genres, including but not limited to "Parking Enforcement", repairs (e.g., "Pothole" or "Tree Issues"), "Trash Pick–Up", "Parks and Recreational Areas", "Noise Complaints", and housing–related urban issues. Users can promote reported issues with "Thanks" votes (similar to "like" in online social media).

We collected information summarized in Table I about all issues pertaining to Albany County reported between January 5th, 2010 and February 10th, 2018 $(2,195)$ using urllib.request[11] and BeautifulSoup[12]. Our dataset includes a list of the SeeClickFix issue report ids. Since every reported issue corresponds to a unique hyperlink of the form https://en.seeclickfix.com/issues/id, our list of ids can be used to harvest detailed information about each reported issue.

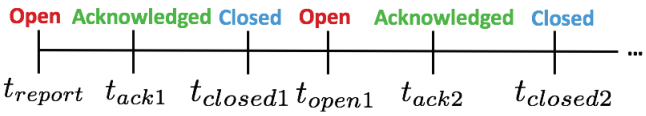


Fig. 2: Illustration of the status timeline of a SeeClickFix reported issue.

Data field	Description
Issue ID	Unique 6 to 7 digit ID of the reported issue.
Title	Title of the reported issue.
Status	"Open", "Acknowledged", or "Closed" (Signifies attention paid by authorities).
Address	Location of the reported issue.
Neighborhood	Neighborhood of the reported issue.
Reporter ID	Unique 4 to 6 digit ID of registered user reporting the urban issue.
Reporter Name	Screen name of registered user reporting the urban issue.
Votes Count	Number of up–votes the reported issue has received.
Thanks Count	Number of "Thanks" the reported issue has received.
Category	The type the reported issue belongs to.
Reported Time	Data and time of the reported issue in UTC$\pm$0.
Reported Via	Medium used to report issue.
Tags	User–defined keywords to simplify the discovery of "similar" reported issues submitted by other users.
Description	Short comment from the user reporting the urban issue.
Q&A	Answers to predefined set of questions.
Comment List	Each comment has (i) a unique 6 to 7 digit ID, (ii) the ID of the user who commented, (iii) status, (iv) comment text, and (v) date and time of comment in UTC$\pm$0.

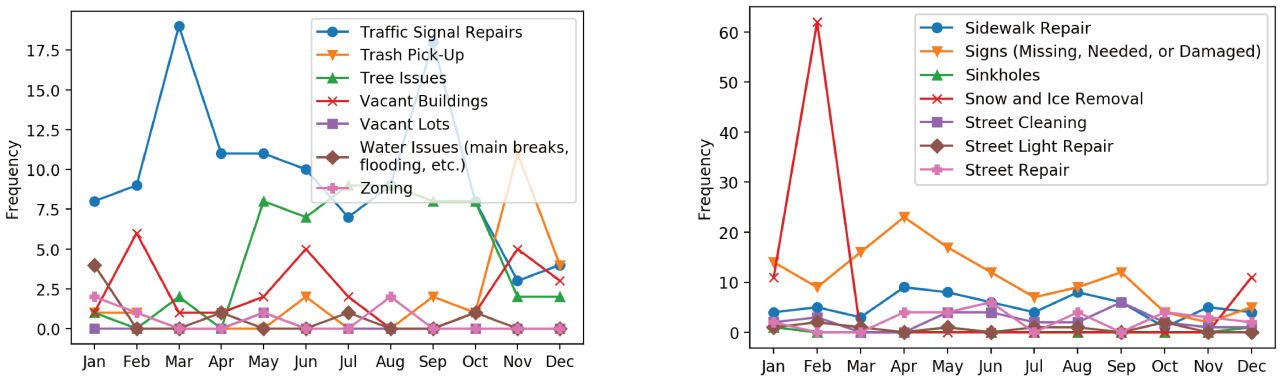
TABLE I: Data for a reported issue in SeeClickFix.

[9]https://blog.seeclickfix.com/seeclickfix-most-engaged-and-most-responsive-local-governments-of-q1-2019-33d20aa60db4

[10]http://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2014/5-year.html

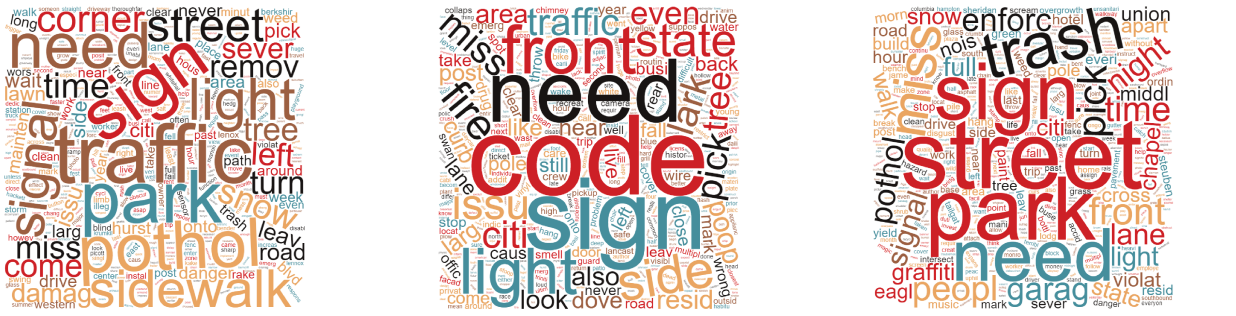
[11]https://docs.python.org/3/library/urllib.request.html

[12]https://www.crummy.com/software/BeautifulSoup/bs4/doc/



(a) (b)

Fig. 3: Number of reported issues per category from May 2013 until February 2018. Reported issues regarding (a) "Traffic Signal Repairs" increase dramatically over March and September, whereas (b) "Snow and Ice Removal" peak in February.



(a) "Arbor Hill" (Mostly Residential) (b) "Center Square" (Mixed) (c) "Downtown" (Mostly Non Residential)

Fig. 4: Word cloud of reported issues for different neighborhood types in our dataset (word size denotes frequency). Pedestrian–related issues (e.g., related to "Traffic Signals" and "Sidewalk Repairs") are more frequent for mostly residential neighborhoods (a), whereas "Code Violations" (e.g., commercial buildings violations) and "Sign Repairs" (possibly as a result of high pedestrian volume) appear more frequently in mixed neighborhoods (b). Issues related to the state of the road surface (e.g., "Potholes") and "Parking Enforcement" are prevalent in all neighborhood types.

Upon initial inspection, reported issues for all categories but "None" have been significantly increasing with time. Furthermore, reported issues have been increasingly assigned to specific categories since their introduction in May 2013, when Albany authorities begun partnering with SeeClickFix, and a verified account associated with a city official was introduced in SeeClickFix. Before this partnership, most reported issues were assigned to the "None" category. Currently, some reported issues are still assigned to the "None" category, however, this is because the reported address dictates the category choices available to users. We also note that the number of reported issues for certain categories exhibit seasonal effects (e.g., "Tree Issues" increase in the months between May and August, whereas "Snow and Ice Removal" requests peak in February, when snow accumulation often reaches its highest for Albany County (see Figures 3a and 3b)). Similarly, certain types of issues seem to prevail in certain neighborhood types (e.g., pedestrian–related issues appear more frequently in mostly residential neighborhoods, whereas code violations are more frequent in mixed neighborhoods (see Figure 4).

In our study, we compute the following statistics from SeeClickFix data at the neighborhood level:

- **Number of Reported Issues**: total number of issues reported within a certain neighborhood.
- **Number of Unique Users**: total number of unique users that have reported at least one issue within a neighborhood. Anonymous reporters are counted as a single user.
- **Number of Non–Anonymous Reporters**: total number of non-anonymous unique users that have reported at least one urban issue within a certain neighborhood.
- **Total Number of Non–Anonymously Reported Issues**: total number of issues reported by non–anonymous users (and corresponding number of reporters).
- **Total Number of Anonymously Reported Issues**: total number of issues reported anonymously (and corresponding number of anonymous users).
- **Total Up–Votes**: total number of up–votes reported issues received within a certain neighborhood.
- **Total "Thanks"**: total number of "Thanks" reported issues received within a certain neighborhood.
- **Response Time**: average duration in seconds between the "Open" and "Acknowledged" (similarly, "Closed") status

of a reported issue. Specifically,

  - **"Open" to First "Acknowledged"**: duration in seconds for a verified account associated with a city official to acknowledge a newly reported issue.
  - **"Open" to First "Closed"**: duration in seconds for the online authority to close a newly reported issue.
  - **"Closed" to $n-$th "Open"**: duration in seconds for a "Closed" issue to be reopened by a user ($n \geq 2$).
  - **$n-$th "Open" to $(n+1)-$th "Acknowledged"**: duration in seconds between the $n-$th reopening and $(n+1)-$th acknowledgement of an issue ($n \geq 2$).
  - **$n-$th "Open" to $(n+1)-$th "Closed"**: duration in seconds between the $n-$th reopening and the $(n+1)-$th closure of a reported issue, where $n \geq 2$.

- **Total Number of Reopenings**: total number of times reported issues are reopened. Although more advanced measures can be considered, in our preliminary analysis we measure reopenings by category and neighborhood.

We use the aforementioned response times to study citizens' reporting behavior as well as frequently recurrent or unresolved issues which would otherwise remain hidden in the data by analyzing reopened issues, their corresponding category and neighborhood they refer to. We underscore that response times are defined with respect to the first "Acknowledged" when the issue is first reported; in other words, the response time is $t_{Response\ Time} = t_{ack_1} - t_{report}$. A total of 953 reported issues have been "Acknowledged" or "Closed" at least once, 69 have been marked as "Open" again at least once, and 9 reported issues have been marked as "Open" at least twice.

Challenges. The neighborhood field is inherently noisy as it depends on the address manually provided by the user who reports an issue. To avoid contaminating our dataset with invalid neighborhood information, we used Google Maps to manually inspect and sanitize the neighborhood field. During this step, a total of 32 neighborhoods were identified, and 609 reported issues were removed from the dataset due to (i) being located outside Albany County, (ii) having a very broad or partially complete address, (iii) being located within the University at Albany campus area, or (iv) being located on the border between two neighborhoods.

Additionally, users reporting an issue in a given neighborhood may not be naturally residing in the neighborhood in which the reported issue refers to. For instance, there is no guarantee that user A, who reported a pothole in the "Center Square" neighborhood actually lives there (i.e., she could be passing by). Thus, a user may report urban issues in multiple neighborhoods, and the set of unique users counted for each neighborhood may be overlapping across neighborhoods.

Finally, the response time for a given reported issue is calculated if and only if an official has either "Acknowledged" or "Closed" an issue. There are 671 reported issues in total in our dataset that have never been "Closed" despite been "Acknowledged" at some point in time by an official. Further research is required to find potential reasons about why some issues remain unaddressed by city officials.

Category	Occurrences
Abandoned Vehicles	36
Albany Housing Authority Issues	20
Animal Control	7
Bike Parking	4
Code Violations	182
Double Pole	3
Fire Hydrant Blocked	27
Fire Hydrant Maintenance	2
Graffiti	16
Illegal Trash	38
Noise Complaints	58
None	122
Other	52
Parking Enforcement	198
Parking Facility Maintenance Assignments (Albany Parking Authority)	2
Parks and Recreational Areas	32
Pavement Markings (Missing, Needed or Damaged)	18
Pothole	79
Property Maintenance (Overgrowth/Grass & Weed Mowing)	70
Recreation and Playground Equipment	12
Sewers, Drainage	1
Sidewalk Repair	62
Signs (Missing, Needed, or Damaged)	125
Sinkholes	3
Snow and Ice Removal	84
Street Cleaning	27
Street Light Repair	9
Street Repair	29
Traffic Signal Repairs	111
Trash Pick-Up	21
Tree Issues	56
Vacant Buildings	27
Vacant Lots	1
Water Issues (main breaks, flooding, etc.)	7
Zoning	6
Total	1,547

TABLE II: Number of reported issues by category.

In the end, 1,586 reported issues remain after manually inspecting and sanitizing the SeeClickFix data.

C. Socioeconomic & Demographic Information from Statistical Atlas

We obtain a total of 40 socioeconomic and demographic statistics (a small subset of which is listed below) at the neighborhood level from Statistical Atlas, which in turn relies upon data from the US Census Bureau.

- **Population Distribution**: the number of entities at the time of survey excluding visitors referred to as *population*, and the ratio of population to total land area excluding water areas referred to as *population density*.
- **Age**: *population* of age cohorts (i.e., 10 to 19 years old, 20 to 29 years old, 30 to 39 years old, and so on).
- **Race and Ethnicity**: *population* of Whites, Blacks, Hispanics, Asians, Mixed and other races.
- **Household Types**: *population* of Family, Married, Single–Female, Single–Male, One person, and other non–family household types.

- **Marital Status**: *population* of marital status (i.e., Never Married, Married, Separated, Divorced, and Widowed) of any residents aged 15 and over.
- **Household income**: *median* household income.
- **Employment Status**: *population* of employed working age (i.e., 35 to 44 years old) men, families with both working parents, families with stay–at–home moms, families with stay–at–home dads, and families with neither parents working.
- **Educational Attainment**: *population* of 25 years of age and older with no high school diploma, with a bachelor's degree, and with a professional or doctorate degree.

Challenges. Some neighborhood boundaries defined in Statistical Atlas do not match the corresponding boundaries defined in SeeClickFix. Specifically, Statistical Atlas does not provide any data for "Beverwyck", "Capitol Hill", "Downtown", "Eagle Hill", "Lincoln Park", "Pastures", "Ten Broeck Triangle", "Washington Park", and "West End". To address this challenge, we replaced missing neighborhoods from Statistical Atlas with the closest census tract or blocks as follows:

- **"Beverwyck"**: census tract 000501, Albany County, NY is used instead of "Beverwyck".
- **"Capitol Hill" and "Downtown"**: census tract 001100, Albany County, NY is used in lieu of "Downtown" and "Capitol Hill" (i.e., issues reported within these neighborhoods are considered to refer to the same geographical area) since each of these neighborhoods either fully or partially overlaps with the tract boundary, respectively.
- **"Eagle Hill"**: census tract 001802, Albany County, NY is used instead of "Eagle Hill".
- **"Lincoln Park"**: there is no appropriate replacement for this neighborhood; thus, all 4 reported issues associated with this neighborhood are removed.
- **"Pastures"**: census block 002500–2, Albany County, NY is used instead of "Pastures".
- **"Ten Broeck Triangle"**: This neighborhood is a sub–neighborhood within the "Arbor Hill" neighborhood. Therefore, issues reported for Ten Broeck Triangle are considered to refer to Arbor Hill; attributes from "Arbor Hill" are used for this neighborhood.
- **"Washington Park"**: There is no appropriate replacement for this area; thus, all 35 reported issues associated with this neighborhood are removed from the dataset.
- **"West End"**: census block 000300–4, Albany County, NY is used instead of "West End".

After re–matching, the number of reported issues from SeeClickFix is further reduced to $1,547$ issues across 28 neighborhoods. Table II shows the final number of reported issues by category, while Figure 5 shows the total number of reported issues by neighborhood in our dataset.

D. Walkability Index, Transit Accessibility, and Bicycle Friendliness Scores from WalkScore

To quantify the ease of which residents can traverse within a neighborhood, we retrieved neighborhoods' Walk, Transit, and Bike scores from WalkScore, a company measuring "walkability" of an address based on the configuration of walking routes to nearby amenities [18]. Although these scores do not reflect all factors that may affect people's transportability (e.g., condition of sidewalks, number of traffic lanes, crime rate, weather), they still provide a useful proxy of street connectivity and accessibility of each neighborhood [18]. Points are awarded using a distance decay function[13], according to which amenities within a 5–minute walk (.25 miles) are given maximum points, and no points are assigned after a 30 minute walk. For reference, each score is defined as follows:

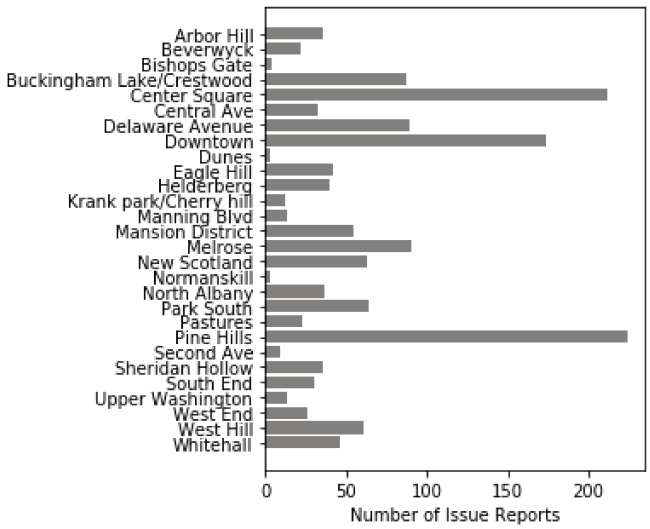


Fig. 5: Number of reported issues by neighborhood.

- **Walk Score**: measures "walkability" on a scale from $0-100$ based on walking routes to destinations such as grocery stores, schools, parks, restaurants, and retail. Small distances yield higher scores, while large distances result in lower scores. This metric estimates "pedestrian friendliness" by analyzing population density and road metrics such as block length and intersection density.
- **Transit Score**: measures how well a location is served by public transit on a scale from $0-100$ by calculating the distance to the closest bus stop on each route, and analyzing the route frequency and type.
- **Bike Score**: estimates on a scale from $0-100$ whether an area is good for biking based on infrastructure (e.g., bike lanes and trails), topography (e.g., hill slope), destinations and road connectivity, and the number of bike commuters.

Challenges. WalkScore relies on Google Maps for the boundary specification of neighborhoods. After aligning neighborhoods from SeeClickFix with neighborhoods from Statistical Atlas, we are left with a 1–to–1 mapping with neighborhood information from WalkScore. Nevertheless, WalkScore does not offer a score for "West End" and "Eagle Hill" neighborhoods. To overcome this challenge, we used the

[13]http://cedeuspubs.geosteiniger.cl/omeka/files/original/b6fa690993d59007784a7a26804d42be.pdf

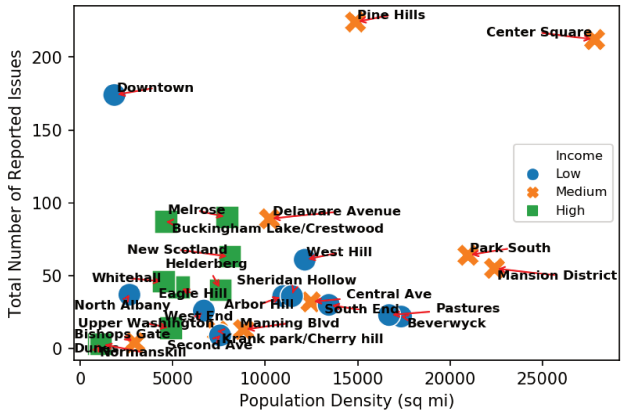


Fig. 6: Total number of reported issues as a function of population density. A slight correlation can be observed. The outliers at the top of the plot (i.e., "Downtown", "Center Square", and "Pine Hills") are all active neighborhoods bustling with office workers, students, commercial stores, and visitors.

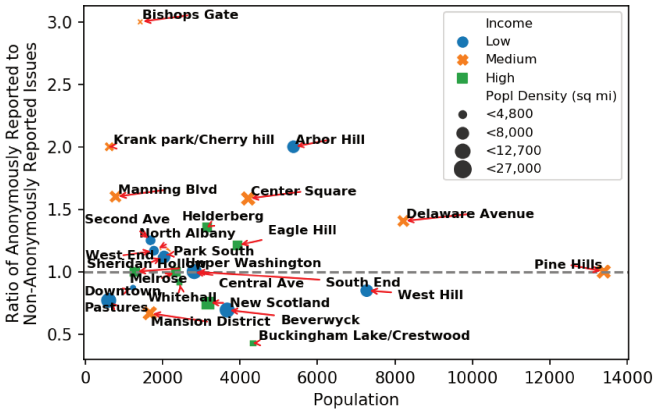


Fig. 7: Ratio of total number of anonymously to non–anonymously reported issues as a function of neighborhood population size.

Walk, Bike and Transit scores of a major street at the center of each neighborhood (i.e., Russell Road for "Eagle Hill" and Watervliet Avenue for "West End").

IV. Results

A. Denser Communities Report More

Here, we examine the link between population density (i.e., number of residents or visitors per square land area) and the aggregate number of reported issues in our dataset. Figure 6 shows a moderate correlation between population density and the number of reported issues. Intuitively, more problems are expected to arise within a highly–populated and dense community as compared to a sparsely populated neighborhood. The "Pine Hills" neighborhood, for instance, hosts a large number of commercial stores and college–aged students. As the neighborhood may attract high foot traffic on a daily basis, it comes as little surprise that "Code Violations", "Traffic Signal Repairs", and "Signs" are some of the most reported issue categories. Conversely, the "Center Square" neighborhood is located near downtown Albany, which is house to many government offices and public buildings (e.g., public libraries and museums). This neighborhood too attracts many visitors, particularly commuters. Thus, "Parking Enforcement" and "Code Violations" are intuitively two of the most reported issue categories. Comparing these two neighborhoods, however, we can conclude that denser neighborhoods may be prone to report more urban issues as "Center Square" reported a similar number of urban issues while it is smaller than "Pine Hills" in terms of both land and population sizes.

B. Report Anonymously or Not?

Since SeeClickFix users can report a urban issue either anonymously or with a registered account, we examine the ratio of anonymously to non–anonymously reported issues as a function of population and average income level, as shown in Figure 7. The clustering of points around where the ratio is close to 1 indicates a delicate balance between the two user types across neighborhoods, whereas the high ratio for the "Bishop's Gate" neighborhood may be an artifact of the small sample for that neighborhood (i.e., only four urban issues have been reported in this neighborhood, three of which anonymously). Nonetheless, since most of the medium–income neighborhoods tend to lie above the $y = 1$ horizontal line, a more detailed analysis is required to verify or disprove a correlation between income and anonymity.

C. Some Users Report More than Others

Next, we plot the ratio of non–anonymously reported issues to unique non–anonymous users as a function of neighborhoods' Walk score (see Figure 8). In general, the number of reported issues is 50% greater than the number of unique users. Focusing on the neighborhoods that do not meet the 1.5 ratio, we notice that the majority of low–income neighborhoods exhibit a 1–to–1 relationship between unique users and reported issues indicating a low rate of reporting per unique user. In comparison, the majority of middle–income neighborhoods are above the 1.5 ratio line. This finding indicates that certain neighborhoods may rely on fewer "regular" observers than others, which may be influenced by the neighborhood's economic status. Future work based on this dataset can examine further the impact of income level on civic participation at the neighborhood level [12], as well as the motivation (if any) of some users to report urban issues more frequently than others.

D. Low–Income Neighborhoods Report Less Despite Higher Walkability

Next, we study the potential effect of walkability, i.e., the level of a neighborhood's "friendliness" to pedestrians in reaching various amenities of interest. One would expect that a high Walk score of a neighborhood would indicate that residents would be more aware of their environment and perhaps report a larger number of urban issues as opposed to neighborhoods with low Walk scores. Figure 9, which illustrates the number of reported issues as a function of the Walk

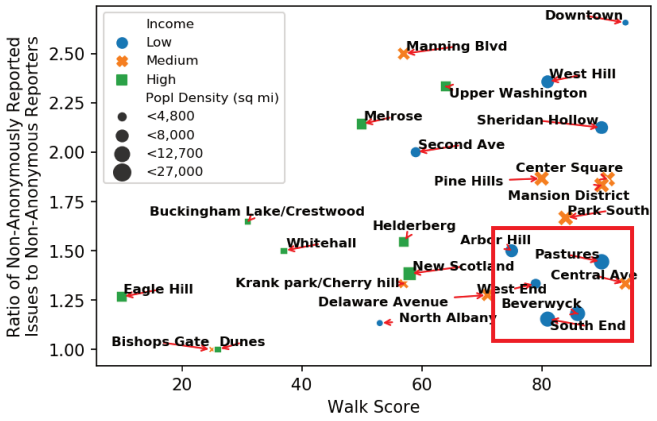


Fig. 8: Number of non–anonymously reported issues divided by the number of unique users as a function of the Walk score.

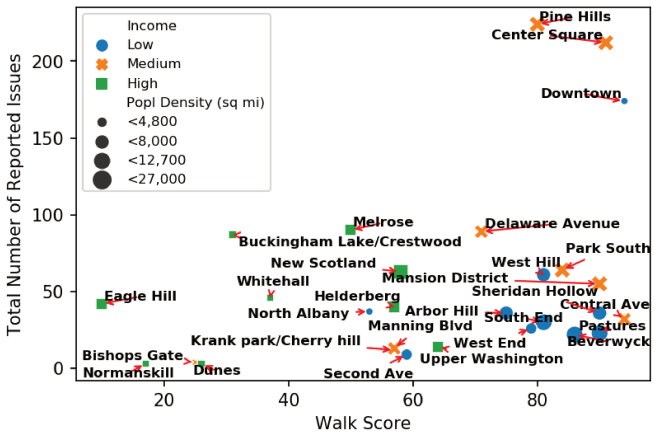


Fig. 9: Number of reported issues as a function of the Walk score of each neighborhood.

score of neighborhoods in our dataset, does not provide strong evidence in support of this hypothesis. "Pine Hills", "Center Square", and "Downtown", all highly walkable neighborhoods as indicated by their Walk scores, indeed demonstrate a high number of reported issues. The high volume of reports for "Downtown" can potentially be attributed to a high volume of visitors (i.e., external observers) in the neighborhood.

An interesting observation is not all highly–walkable neighborhoods report more urban issues. In fact, low–income neighborhoods, despite their high Walk scores, do not seem to report a lot of urban issues. Specifically, Figure 9 shows that low–income neighborhoods cluster at the bottom right of the plot, while medium–income neighborhoods are positioned near the middle as well as the top right of the plot. Future research can further examine the holistic meaning of "walkability" and other factors that may influence walkability in high– and low–income neighborhoods.

E. Income Does Not Dictate Median Response Time

Herein, we examine potential bias in response to reported issues by officials for low–, medium– and high–income neighborhoods. Figure 10 shows that the bias hypothesis cannot be verified as the median response time by officials does not fluctuate based on the average income level of neighborhoods (i.e., no prejudice against certain neighborhoods is observed).

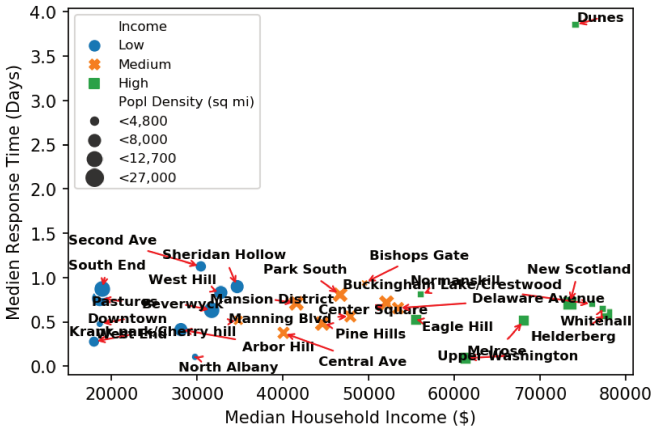


Fig. 10: Response time by neighborhood as a function of median income.

F. Reopened Issues: "Closed" but not Fixed?

Finally, we analyze the total number of reported issues that have been reopened by category (see Figure 11) and by neighborhood (see Figure 12). We observe that "Traffic Signal Repairs" and "Code Violations" are among the top reopened reported issues in our dataset. Further examining the distribution of reopenings by neighborhood, we observe that "Pine Hills" and "New Scotland" neighborhoods feature a high number of reopenings of issues regarding "Traffic Signal Repairs" (see Figure 12). Although further research is required to explain the emergence of such patterns, our observation demonstrates the benefit of using this dataset to assist in identifying neighborhoods in which the same problems are reported multiple times, which in turn may indicate frequently occurring issues, engaged citizens, high foot traffic, lack of city's response, or failure of implemented solution.

V. Open Research Questions

The increasing popularity of civic engagement platforms suggests that they could act as a powerful research tool for understanding their growing use and related implications. With this in mind, we list few ambitious questions that we believe the research community may wish to explore in future work:

1) Why do certain neighborhoods report urban issues more than others? What makes those neighborhoods uniquely more engaged than others?
2) What factors may impede low–income neighborhoods from reporting neighborhood issues?
3) How closely do the urban issues reported online match the issues that actually exist off–line? Could the matching rate differ by neighborhood?
4) What type of urban issues are reopened in certain neighborhoods and why?

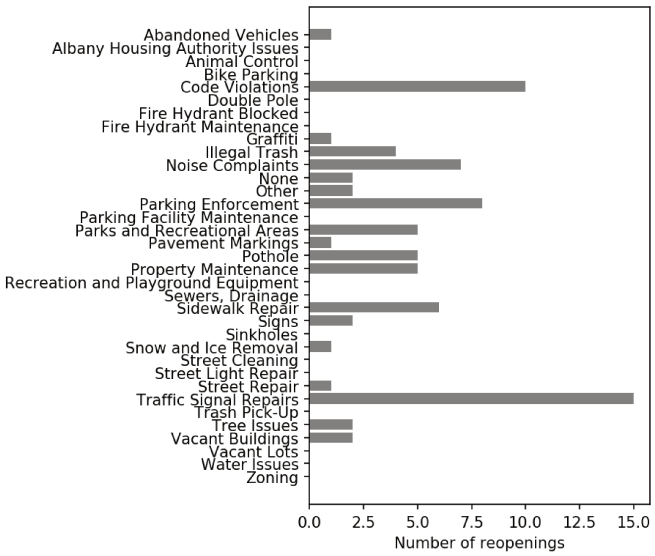


Fig. 11: Number of reopenings by category.

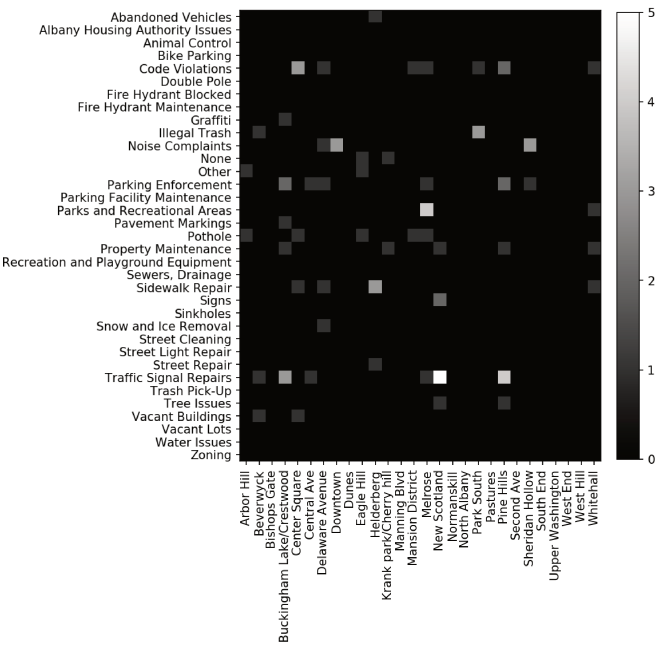


Fig. 12: Number of reopenings by neighborhood and category.

Prior research provides several conceptual frameworks (e.g., community infrastructure, collective efficacy) to help researchers explore how to answer some of these questions [19], [20]. Although additional qualitative insights are needed to fully answer the questions above, our study provides the foundational context and hypotheses to move forward with addressing such ambitious questions.

VI. Conclusion

In this work, we we have illustrated how the environment in which people reside may influence their issue–reporting behavior. Future research can build upon our findings and a number of other variables available in our dataset (e.g., racial/ethnic background, educational attainment, Transit/Bike scores), which we open up completely to the public, to reveal additional correlations that can help explain why certain neighborhoods show high or low civic engagement.

References

[1] I. Mergel, "Distributed Democracy: SeeClickFix.com for Crowdsourced Issue Reporting," 2012.

[2] S. F. King and P. Brown, "Fix My Street or Else: Using the Internet to Voice Local Public Service Concerns," in *1st International Conference on Theory and Practice of Electronic Governance*, 2007, pp. 72–80.

[3] S. A. Chun, S. Shulman, R. Sandoval, and E. Hovy, "Government 2.0: Making Connections Between Citizens, Data and Government," *Info. Pol.*, vol. 15, no. 1,2, pp. 1–9, April 2010.

[4] J. A. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, and M. B. Srivastava, "Participatory Sensing," 2006.

[5] P. Dutta, P. M. Aoki, N. Kumar, A. Mainwaring, C. Myers, W. Willett, and A. Woodruff, "Common Sense: Participatory Urban Sensing Using a Network of Handheld Air Quality Monitors," in *7th ACM Conference on Embedded Networked Sensor Systems*, 2009, pp. 349–350.

[6] R. K. Ganti, N. Pham, H. Ahmadi, S. Nangia, and T. F. Abdelzaher, "GreenGPS: A Participatory Sensing Fuel–Efficient Maps Application," in *8th International Conference on Mobile Systems, Applications, and Services*, 2010, pp. 151–164.

[7] P. Zhou, Y. Zheng, and M. Li, "How Long to Wait? Predicting Bus Arrival Time with Mobile Phone Based Participatory Sensing," *IEEE Transactions on Mobile Computing*, vol. 13, no. 6, pp. 1228–1241, 2014.

[8] A. L. Kavanaugh, E. A. Fox, S. D. Sheetz, S. Yang, L. T. Li, D. J. Shoemaker, A. Natsev, and L. Xie, "Social Media Use by Government: From the Routine to the Critical," *Government Information Quarterly*, vol. 29, no. 4, pp. 480–491, 2012.

[9] D. C. Brabham, "A Model for Leveraging Online Communities," *The Participatory Cultures Handbook*, vol. 120, 2012.

[10] C. J. Tolbert and K. Mossberger, "The effects of e-government on trust and confidence in government," *Public administration review*, vol. 66, no. 3, pp. 354–369, 2006.

[11] R. M. Emerson, "Social exchange theory," *Annual review of sociology*, vol. 2, no. 1, pp. 335–362, 1976.

[12] M. L. Ohmer, "How Theory and Research Inform Citizen Participation in Poor Communities: The Ecological Perspective and Theories on Self– and Collective Efficacy and Sense Community," *Journal of Human Behavior in the Social Environment*, vol. 20, pp. 1–19, 2010.

[13] H. Lelieveldt, K. Dekker, B. Volker, and R. Torenvlied, "Civic Organizations as Political Actors: Mapping and Predicting the Involvement of Civic Organizations in Neighborhood Problem–Solving and Coproduction," *Urban Affairs Review*, vol. 45, pp. 3–24, September 2009.

[14] D. T. O'Brien, E. Gordon, and J. Baldwin, "Caring About the Community, Counteracting Disorder: 311 Reports of Public Issues as Expressions of Territoriality," *Journal of Environment Psychology*, vol. 40, pp. 320–330, 2014.

[15] J. S. Coleman, "Social capital in the creation of human capital," *American journal of sociology*, vol. 94, pp. S95–S120, 1988.

[16] N. Lin, "Building a network theory of social capital," in *Social capital*. Routledge, 2017, pp. 3–28.

[17] J. Ju, L. Liu, and Y. Feng, "Understanding continuous citizen participation on a green commuting platform: The roles of public value and private value," in *Proceedings of the 52nd Hawaii International Conference on System Sciences*, 2019.

[18] L. J. Carr, S. I. Dunsiger, and B. H. Marcus, "Walk Score as a Global Estimate of Neighborhood walkability," *American Journal of Preventive Medicine*, vol. 39, no. 5, pp. 460–463, November 2010.

[19] J. T. Carbone and S. E. McMillin, "Neighborhood Collective Efficacy and Collective Action: The Role of Civic Engagement," *Journal of Community Psychology*, 2018.

[20] Y.-C. Kim and S. J. Ball-Rokeach, "Civic Engagement from a Communication Infrastructure Perspective," *Communication Theory*, vol. 16, no. 2, pp. 173–197, 2006.