Research Article

ICONARP
International Journal of Architecture and Planning
Received: 22.03.2021 Accepted: 31.08.2021
Volume 9, Issue 2/ Published: 21.12.2021
DOI: 10.15320/ICONARP.2021.181 E- ISSN:2147-380

ICONARP

Data Mining the City: User Demands through Social Media

Hülya Soydaş Çakır[1], Vecdi Emre Levent[2]

[1]Asst. Prof. Dr., Faculty of Engineering and Architecture, Fenerbahçe University, İstanbul, Turkey. (Principal contact for editorial correspondence), Email: hulya.soydas@fbu.edu.tr

[2]Dr., Faculty of Engineering and Architecture, Fenerbahçe University, İstanbul, Turkey. Email: emre.levent@fbu.edu.tr

Abstract

Purpose

Information technologies are commonly used in architectural and urban design. The use of these technologies providing support at every stage of the design opens up different perspectives for designers and users. The aim of the study is to obtain user demands for green spaces of a specific district by mining data through social media and to detect the actual green spaces of the same district using applications developed for this purpose. User demands for design decisions and applications of green spaces and the current situation of the study area are evaluated.

Design/Methodology/Approach

The research is firstly realized through social media, and data obtained from Twitter is analysed in order to evaluate user demands for parks and green spaces of Ataşehir district in İstanbul City. Secondly, all green areas in the same district are detected by using digital maps. Two applications are specifically designed for this research; Tweet Grabber is used for user sentiment analysis on social media and Map Grabber is processed for extraction of green spaces via maps. The total area of the green spaces is compared with the desired area of open and green spaces per user.

Findings

The user demands and thoughts obtained in the study about the green spaces of the district are compatible with the actual situation of green spaces. It is observed that the users are mostly dissatisfied with the adequacy of green spaces. Designers, politicians, municipalities and all stakeholders can benefit from the obtained user expectations and feedback. Interpreting user demands by mining data through social media enables user participation in design decisions. This research method can be supportive and adaptive in related issues of design for the cities, enabling user participation in architectural and urban design.

Research Limitations/Implications

Parks and green spaces of Ataşehir district of İstanbul are taken as a case study. Twitter is chosen for mining of data in social media based on parameters such as keywords and location.

Social/Practical Implications

The impact and support of users in design decisions can be clearly demonstrated by advanced information technologies. Mining data through social media and developed applications will contribute to design decisions and policies for architectural and urban spaces.

Originality/Value

Tweet Grabber and Map Grabber applications are developed for this research in order to get text based and image based data. The research includes a unique case study for mining data through social media on a specific design issue and target location.

Keywords: *Architecture, city, data mining, design, social media.*

INTRODUCTION

Dissemination of information is faster and more effective today than in the past decades. More data can be accessed more quickly in each business sector. Obtaining and interpreting user data in design issues have been carried out with traditional methods for centuries. Nowadays, the process of obtaining and evaluating this data is changing with the opportunities provided by the rapidly increasing acceleration of information technologies. Studies using data mining method have started to play a critical role in identifying different problems in design issues and developing solutions under this topic. The use of user sentiment analysis and map processing techniques has been revealed to allow a researcher in design field to process data that is far too big for manual processing. The analysis of urban data helps to get information about the city and opinions of the citizens for planning and decision making (Psyllidis et al., 2015). Resch et al. (2016) carried out a study on evaluating emotions of the citizens by analysing data from social media and they conclude that extracting emotions from social media can remark planning aspects of cities; and the emotions, thoughts and expectations about the city detected by tweets can be utilized in urban planning process. Technology enables the citizen participation in public planning projects through the medium of Web (Brabham, 2009). Tasse and Hong (2014) inform in their study that social media data can help planners for future projects, can be useful for measuring the pulse of residents and for their integration to new cities. Salesses et al. (2013) indicating that available data for urban perception is limited, used a method in their study based on online image ratings to quantify how users perceive the cities. They pointed at the future development of different techniques to explore how urban issues are perceived. As Seltzer and Mahmoudi (2013) inform in their study, citizen involvement is the result of multiple techniques in planning processes enabling to make better decisions for liveable communities with equal opportunities.

In this study, it is pointed out that access to user demands, criticisms and suggestions in architecture and urban design can be realized through social media. The data obtained by data mining method from a district is evaluated in this research. The study is proposed to support user-oriented design in urban environment. In the research, Istanbul was selected as central province and Ataşehir as an example district in order to reach user data. This study has been focused on the determination of the current status of public green areas and the user comments for these spaces. The research includes specially developed software that are used for mining data through social media and visual map analysis to collect citizen insights on planning decisions about green areas. In this study green space encompasses parks, gardens, forests and groves. The demands of the citizens about all green spaces on the determined location are evaluated. This study discusses how the methods designed for this research can be adapted to different research

DOI: 10.15320/ICONARP.2021.181

in the field of design. This kind of approach to similar research is unprecedented with its specific software deliberately developed for this case study and with the determined location.

DATA MINING IN ARCHITECTURAL AND URBAN DESIGN

The rapid development of information technologies brings about changes in building and urban scale. With the concept of smart buildings and smart cities, sustainable designs are handled in a user-oriented manner and studies are carried out to increase service quality and efficiency. The emergence of smart cities and big data concepts provide new potential platforms to resolve various urban diseases (Pan et al., 2016). The analysis of big data also provides predictions of urban life and allows alternative views for urban development by supporting everyday life and decision-making processes (Kitchin, 2014). There are research to determine the current status and identity of the city for efficient planning in urban development and correct operated design processes. Different scientific methods have been applied in these research. Chang et al. (2017) explored the different social areas of the city of Zurich, including parks, gardens, playgrounds and squares using data mining technologies. They pointed out that the interpretation of the available data with the methods they used could guide the designers in determining the urban identity characteristics for social areas. A similar study was carried out by Chang et al. (2018) for the city of Taipei where data mining technologies were used to determine both the socio-economic behaviors of users and the current status of social areas. Based on the results obtained, an urban design workshop was organized, and the views and comments of citizens in the urban design decisions were taken over the web platform and were incorporated in designs. Access to more responsive and liveable urban ideals has been provided with their study. In a study conducted by Valls et al. (2018), it was concluded that it is possible to obtain valuable data and information to determine the different uses and architectural requirements of urban space, but it is emphasized that this data may be challenging to retrieve, structure, analyze and visualize. As stated in the same research, social media applications can be used as a tool to evaluate user responses in public participation processes before final project, as well as obtaining data for the initial definition of the project and feedback after the completion. Münster et al. (2017) pointed out that the detailed planning process does not constitute a guarantee for the broad acceptance of an envisioned urban project. Therefore, participation of inhabitants in urban planning processes, the digital environments and digitally supported approaches are important for this purpose. As it is mentioned in their research; social media is a virtual and two-way communication channel for user participation in urban planning.

ICONARP – Volume 9, Issue 2 / Published: 21.12.2021

Data mining and methods of use

The interpretation of a large number of data produced everyday can no longer be processed manually. Therefore, data mining has become a very necessary field of study. Data mining is very useful for analysing situations where the relationship between input and output is complex and/or when there is a large amount of data. Data mining is commonly characterized by volume, velocity, variety, veracity and value concepts (N. Khan et al., 2014; Yang et al., 2017; Zikopoulos et al., 2012). Volume that is the magnitude of large-scale datasets (Rao et al., 2018) increases every day and the size of the data set increases as well. Velocity refers to the speed of change in data as social media posts (Hashem et al., 2015). Various data is stored from many resources, and different types of data can be used for information. Big data can be structured, unstructured or semi-structured data. Veracity represents the quality of data as it is reliable and generated correctly. The data should be convenient for the purpose and be up to date. Many benefits can be obtained by analysing and processing the big data referring to its value characteristic (Elragal, 2014; Younas, 2019).

Data mining algorithms are generally divided into two categories: Classification and Clustering. Classification algorithms are useful in predicting the label of a data that will come without tag information by training in a dataset containing tag information. Clustering algorithms are used to divide the dataset by the number of selected clusters without any tag information in the dataset obtained. The aim of this process is to keep the distance of the clusters to each other to the maximum and to keep the distance between the elements in the cluster to a minimum. Both main methods have many sub-algorithms. Classification algorithms include artificial neural networks, Support Vector Machines (SVM), Decision Tree (Lessmann et al., 2015; Nguyen et al., 2014; Wong, 2015). All these algorithms generate a model by training a training set. The test data is fed to the model, and it is expected that the data will be classified correctly. Clustering algorithms include k-means, fuzzy c-means, density-based clustering algorithms (Allahyari et al., 2017).

The term data mining is commonly encountered in various sectors such as engineering, health, education, marketing and banking. The use of data mining is applied in order to process data having great importance in order to reach knowledge in the collected information. In our daily lives and in many sectors, the storage and analysis of the cumulating data becomes more difficult. However, the processing and interpretation of data becomes inevitable in the process of producing the required information. Data mining is broadly defined as the process of discovering interesting patterns and information from large amounts of data. These data sources can be databases, data warehouses, Web, other information repositories or data that is dynamically transferred to the system (Han et al., 2012). The ability to analyse data in the relevant field provides a significant advantage for those offering products and

DOI: 10.15320/ICONARP.2021.181

services in that sector. Therefore, in the field of architecture and urban design, various studies are held by using data mining approach. When the importance of user types, characteristics, needs and demands at the design stage is taken into consideration, the analysis of available data will be valuable in design processes. The studies will especially prove to be beneficial in public open spaces where social life is shaped and social habits and preferences emerge.

Mining of data through social media

Obtaining and interpreting data in a medium where users can express themselves, thoughts and feelings guides researchers and planners. The undeniable power and rise of social media have given importance to scientific studies in this field. Social media has become a platform where real-time user views are shared extensively. It also constitutes an important source for obtaining positive or negative perceptions and comments in many cases. User demands in cities are shaped by a wide variety of users.

The expansion of urban big data has an important role in the development of cities. Urban big data types are summarized under different topics. The type of data that includes examples such as social media, web usage, GPS, online social networks is called user-generated content (Thakuriah et al., 2017). Urban-scale research through social media is highly valuable in terms of access to large audiences. The positive contributions to decision-making processes in planning and design stages are emphasized in scientific research conducted through social media. In one example of these studies; Chen et al. (2017) used TripAdvisor social media application to examine how the city is perceived by tourists. Positive and negative comments were analysed and the issues that should be focused on planning and design in terms of parking problems were determined. They pointed out that further studies carried out via social media could include matters such as land use, transportation, open spaces, health, education and others. In another study conducted by Chen et al. (2016) for the city of Boston, data mining was carried out through different social media sources to uncover problems such as poorly or improperly used areas. Residents who share information about their environment through social media can complete or even change the information measured by physical sensors. Human perception provided by social media has the potential to support smart city initiatives (Doran et al., 2015). Mueller et al. (2018) define design feedback of the inhabitants as an essential way for a responsive city by integrating their ideas and wishes in design processes. Thus, citizen participation will contribute to transdisciplinary research studies. The intensive and widespread use of social media by the inhabitants can make it easier to obtain large-scale data on many different areas of the city. Planners and politicians can reach the opinions of the citizens on various issues about architectural and urban design. On the other hand user participation has its challenges as there is always the risk of criticism.

METHODS

This research is based on two applications for the purpose of the study. Software called Tweet Grabber is developed to obtain a dataset of tweets from Ataşehir district in İstanbul containing keywords related to selected keywords and location. A selected algorithm is applied to detect comments about green spaces in the district by using the dataset based on user tweets. The green spaces in the district are calculated by the other developed software called Map Grabber. This software is used to detect all green areas automatically by using digital maps. The result of this detection is used to calculate the green spaces per inhabitant and to interpret the data obtained. In Figure 1, the data collection and processing methodology of the system are demonstrated.

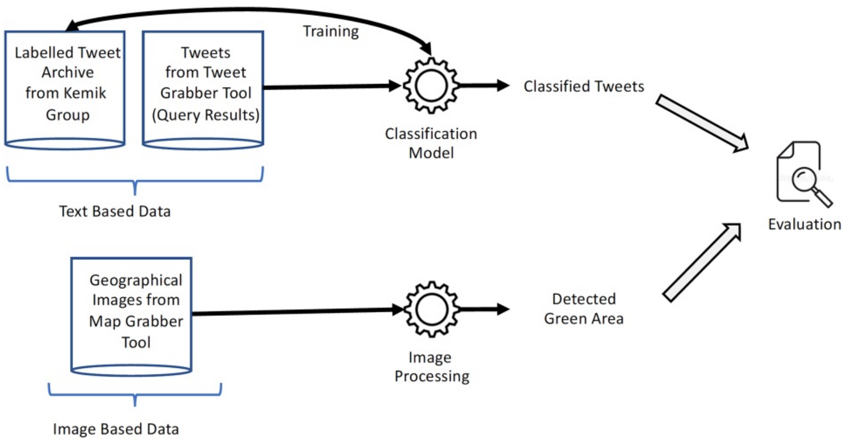

Figure 1. Data collection and processing methodology

The results of tweets classified in the study are compared with the status of the green areas detected within the district. It is aimed to determine how the existing green spaces and services are evaluated by the users, what their wishes and expectations are, and what their suggestions and needs are for future projects.

Data collection on social media

Data is needed to analyse people's emotions through social media. Twitter was selected in this study as the source of social media data. Twitter has a growing popularity for leading user generated content (Kotzias et al., 2016) and acts as a valuable method to gather user viewpoints (Hasib et al., 2021). Twitter enables interactions between the users and institutions and the feedback help organizations to improve their services or products as well (Salur & Aydin, 2020). It is possible to search and collect tweets with various keywords on Twitter. Twitter allows access to their data with some limitations through its API (Valls et al., 2018). However, performing this process manually may take some time. For this reason, software called Tweet Grabber has been

DOI: 10.15320/ICONARP.2021.181

developed that automatically saves tweets by searching with parameters such as keywords, location where the tweet was sent. A library called "tweetinvi" was used in the software infrastructure. There are similar studies in the literature (Chatterjee & Perrizo, 2015; Kobayashi et al., 2016). In order to use this library, it is necessary to take developer permission from Twitter and enter the developer ID codes given by Twitter into the library. The screenshot of the developed tool is given in Figure 2. This application is developed with C# language using the Tweet capture API provided by Twitter. With the developed interface, various parameters can be received from the user and fed to the Twitter API. Twitter has query limits because the Twitter API provides tweet capture. Within these limits (300 Tweets in 15 Minutes), multiple queries were made at different times to create the necessary tweet database.

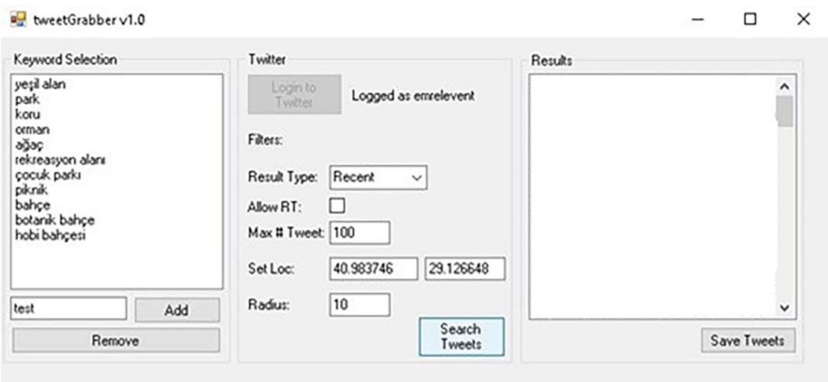

Figure 2. Tweet Grabber application

Since the study was in Turkey/İstanbul/Ataşehir location, Turkish keywords were used to search data to be received via Twitter. The keywords used are given below.

- Green Area (Yeşil Alan)
- Park (Park)
- Grove (Koru)
- Forest (Orman)
- Tree (Ağaç)
- Recreation Area (Rekreasyon Alanı)
- Child Park (Çocuk Parkı)
- Picnic (Piknik)
- Garden (Bahçe)
- Botanic Garden (Botanik Bahçe)
- Hobby Garden (Hobi Bahçesi)

The software collects keywords by scanning them on Twitter. At the end of the process, related tweets were collected in a text file.

Finally, the tweet archive prepared by the "Kemik" study group at Yıldız Technical University was taken to be used as training data in the sentiment analysis (Amasyali et al., 2018). In this database, there are tweets containing the name of a commercial company from Turkey the database of which has been labelled by the Kemik research group. This archive contains approximately 17k tweets and tag information for each tweet. The labels that were used were positive, neutral and negative.

Map extraction

Detecting green areas in urban space is an important problem. This process can be done in several ways. The first is to obtain the data of the relevant municipalities. The second can be manually identified green areas through maps. Finally, the maps can be loaded into an application, and green areas can be detected automatically.

There are many studies in the literature for automated green field detection. These are studies made with satellite map images. These studies basically take the images of online map providers such as Google Maps, Google Earth and OpenStreetMap and process them with various image processing techniques such as vegetation extraction, green area detection and water resource detection. Google and OpenStreetMap allow data to be retrieved and processed from their maps as long as the data is not used for commercial purposes (Google, 2021; OpenStreetMap, 2021). OpenStreetMap contains vectors; however, it is not possible to mark each green area separately. For example, a green area in the garden of an apartment will not be identified. In this respect, the system success performance of working with marked vectors will decrease. In a study with images taken from Google Earth (Almeer, 2012), vegetation extraction is performed using the Back Propagation Neural Network algorithm. Pre-recorded images are fed to the algorithm for training, so that the vegetation is learned. It works successfully in places such as deserts, cities and highways. In another study conducted on Google Maps (Hegadi & Sangolli, 2011) segmentation was made according to colour characteristics; for example a coastal view distinguishes between sea and green fields. In the study by using the deep convolution neural network algorithm (Kaiser et al., 2017), it is possible to distinguish buildings and streets on maps obtained through OpenStreetMap. In the study of street level foliage analysis (Li et al., 2015), street images were obtained via Google Street View. The green areas were determined according to the colour filtering on the street pictures taken.

In this study, the Map Grabber software is developed in C# like the Tweet Grabber software. The purpose of this software is to detect all green areas in Ataşehir district borders. For this purpose, firstly all the borders of Ataşehir location were drawn as a polygon. With the developed software, all maps in the given borders were collected. Figure 3 shows the software from which the maps were obtained. As a result of this process there were 124 map images in total.

DOI: 10.15320/ICONARP.2021.181

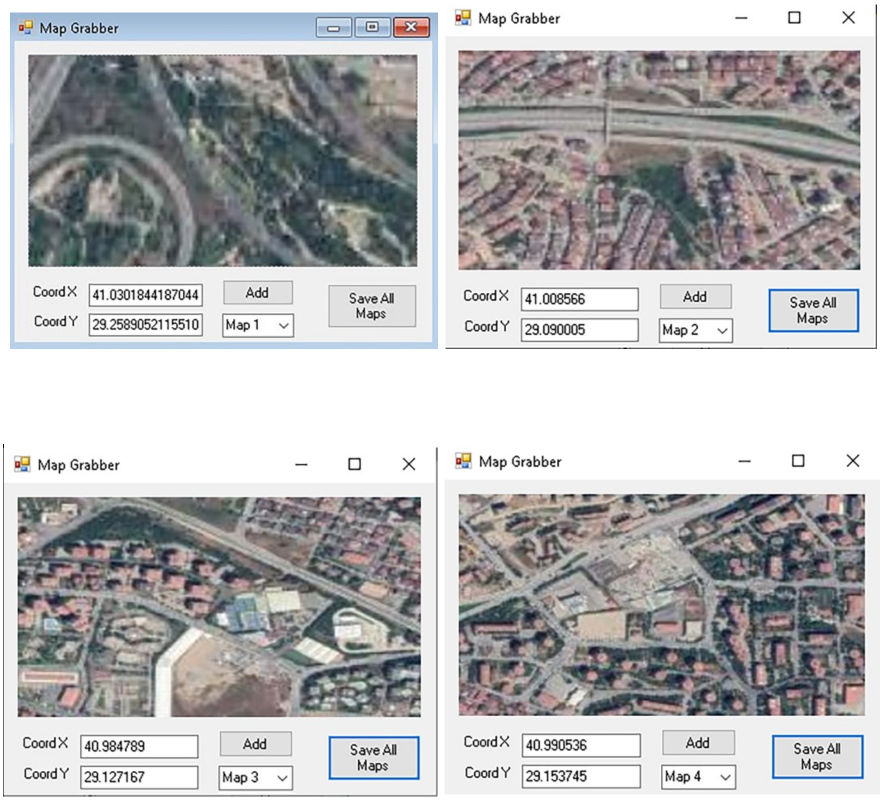

Figure 3. Map Grabber application

Sentiment and map analysis methods

Sentiment analysis method is one of several methods used to draw meaningful results from the data obtained through social media in order to improve product and service quality. A study was carried out to analyse emotions and classify them in categories of news, politics and culture by using Bayesian classification algorithm (Baykara & Gürtürk, 2017; A. Khan et al., 2015; Pang & Lee, 2008; Santos & Gatti, 2014). In another study, word, semantic and character-based analysis was conducted for Turkish sentiment analysis (Amasyali et al., 2018).

In this study, a data set was created by searching related keywords on Twitter. Each of the elements in this dataset should be classified as positive, negative or neutral. A tweet archive of approximately 17k was used for the training process, previously tagged by other researchers (Amasyali et al., 2018). There are 4500 positive, 6800 negative and 5800 neutral tweets in the dataset used.

Six different approaches have been applied for the classification of the dataset. These approaches are listed below:

• Bag of Words (BoW): A matrix is created for all words in all documents. For each document, the number of words (TF) is recorded. It is then normalized with the reverse document frequency (TF-IDF). The extracted values are classified by machine learning algorithms such as SVM. In literature (Joulin et al., 2016; Schmidt & Wiegand, 2017), this algorithm is being used for hate recognition and author classification has been conducted.

- N-Gram Bag: Indicates subtracting the repeat rate in a given sequence. N refers to the size of the sequence to be found in the text. A matrix is created based on the number of sequence sizes selected. The outputs are classified by a machine learning algorithm as in the Bag of Words method.
- Word Vectors: The approach groups synonymous words and for this, it tries to find a space where the coordinates of the words with close meaning are close. A tool called Fasttext provides ready-made training vectors. With these vectors, word vectors are obtained with training examples. With Word2vec technique, the words in the text are expressed as a vector. The sample to be tested is matched to the closest group by looking at the spatial distance.
- Long Short-Term Memory: In addition to whether the words are in the text or not, how words are arranged is also important. The sequence between words can be modelled with Long Short Term Memory algorithm.
- Character Based Convolution Neural Networks: Includes embedding characters instead of words. The dimensions of word vectors can be very large. However, character vectors are quite small compared to word vectors. The algorithm generates the results with CNN algorithm by hovering filters over the generated matrix.

"Character Based Convolutional Neural Networks CB (CB-CNN) algorithm was found to be a more successful classification method in the literature (Amasyali et al., 2018). With the selected (CB-CNN) algorithm, a model was created with 17k training data. The model created at this stage was classified by feeding tweets from Ataşehir location. After the sentiment analysis, the amount of green space has been extracted from the maps in order to compare with the current situation.
Map analysis is an important issue in green field extraction. Green areas were determined from the raw images taken. An example of raw image is shown in Figure 4.

Figure 4. Captured raw image

DOI: 10.15320/ICONARP.2021.181

The raw image is first converted from the RGB colour space to the LAB colour space. The image is then segmented by colour by using K-Means algorithm. In the resulting clusters, clusters with green tones are selected, the remaining clusters are coloured with black. The processed image is shown in Figure 5.

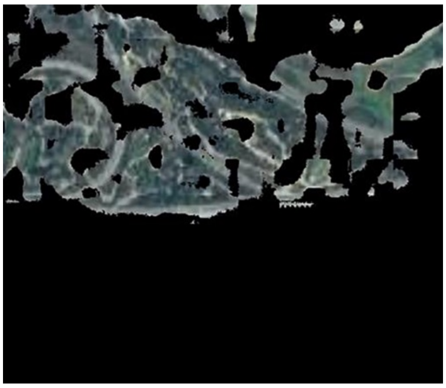

Figure 5. Processed image

The number of pixels remaining in the processed image indicates the amount of green space. Based on the distance view point on the map, the total area is calculated based on the amount of area each pixel represents.

RESEARCH FINDINGS

Current situation analysis

The green space in Ataşehir district was determined by using the image processing techniques. The total green area in the region is calculated to be 511k square meters. As reported by TUIK (Turkey Statistical Institute), Ataşehir district has a total population of 416,318 in 2018 (TUIK, 2019). The District Municipality is updating this total as 422,513. District area is totally 25.84 square kilometres (Ataşehir Municipality, 2020). The district appears as a region where construction activities and transportation density increase rapidly. The region is a developing part of the city with the potential demands of the residents.

It is stated that the children's playground, park, square, district sports area, botanical park, recreation area and recreation infrastructure areas should be at least 10 square meters per person according to the plans made by the Ministry of Environment and Urbanization within the borders of district under the title of Open and Green Areas. This size was determined as 5 square meters per person for plans within the city borders including infrastructure spaces such as zoo, urban forest, afforestation area, exhibition, fairground/festival area and hippodrome. The same minimum area sizes are specified in the population groups of 0-501.000 and over (Ministry of Environment and Urbanization, 2019). In this research, open and green space area for each person is found to be less than 2 square meters within the district. The ratio of open and green areas against the total district area is below 2%. This rate is far below the percentage of public green space in many world cities (e.g.

ICONARP – Volume 9, Issue 2 / Published: 21.12.2021

Amsterdam 13.0%, Cape Town 24.0%, London 33.0%, Hong Kong 40.0%) (World Cities Culture Forum, 2019).

Interpretation of Research Data

In the study, the tweets of the users from Ataşehir district about the current status of green spaces were classified and fed to the model. There are 2969 tweets that are fed to the model. 72% of these tweets were negative, 18% positive and 10% neutral (Figure 6).

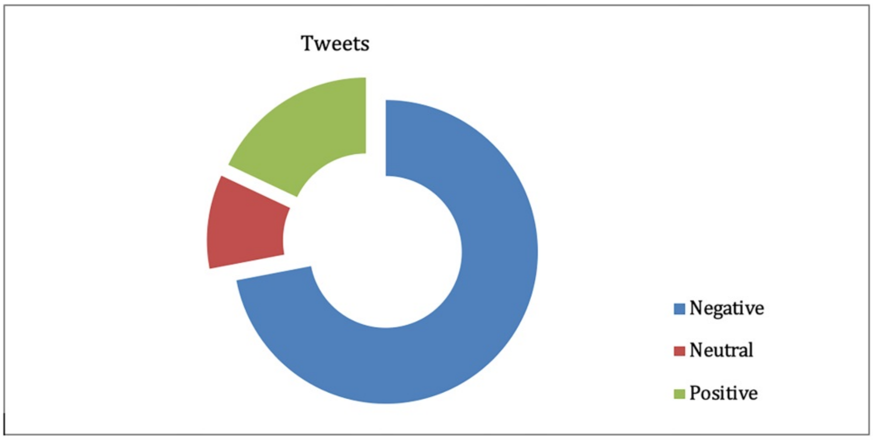

Figure 6. Interpretation of tweets

The frequencies of words in tweets are summarized in two groups. In the first group, there are keywords used for searching data (Table 1) and in the second group the most used adjectives, pronouns, punctuations and other items are listed (Table 2).

Table 1. Frequency of keywords in tweets

No	Word	Frequency
1	picnic	534
2	park	475
3	area	416
4	green	402
5	garden	399
6	grove	384
7	tree	296
8	child	166
9	forest	142

Almost all searched keywords were encountered in tweets. The words "picnic", "park" and "green/area" were mostly used by the residents. A significant majority mentioned the concepts of "garden", "grove" and "tree." In the second group, it is remarkable that exclamation and question marks are frequently used in the statements examined. It is observed that users are very responsive and questioning about parks and green areas. Words such as "a/one", "none", "more", "much/many", "but" exist widely in tweets. Users provided remarkable explanations and warnings about green areas within the district.

DOI: 10.15320/ICONARP.2021.181

User opinions about green areas are predominantly negative with a large quantity of 72%. In negative tweets, it is emphasized that green spaces are not cared enough, construction facilities are constantly increasing and conscious urbanization is not taken into consideration. Many tweets indicate that the green areas are inadequate and poor; requests for improvement are conveyed and criticism is included in planning and goal setting issues. Users with negative opinions have a critical approach to review environmental and urban planning policies. The needs for parks and green spaces are also expressed consistently. The residents state in many tweets that while urbanization increase in the form of concrete construction, green areas are gradually decreasing. They criticized the authorities for failing to take precautions in some cases. Users complained about the lack of parks and green spaces in many locations.

Table 2**.** Frequency of other items in tweets

No	Word	Frequency
1	a/one	792
2	and	455
3	!	439
4	?	420
5	this	374
6	you	222
7	what	213
8	for	190
9	none	166
10	more	138
11	much/many	137
12	every	129
13	but	119
14	beside	102
15	with	96

When the content of positive tweets (18%) in the research is evaluated, it is found that individual areas and landscapes, some of the park samples and green space projects are appreciated. Some of the users told their personal opinions about sample projects and reported their likes. There are no positive opinions or notifications on green space adequacy, number of parks and the investment/planning policies throughout the district.

The tweets reviewed contain only 10% neutral comments and opinions by users for the current situation. These users shared their opinions on social media under the title of green spaces, but no positive or negative opinions were found related to the study. However, the majority of neutral interpretations have wishes for the extension of green areas and parks. Although they do not contain negative expressions, those tweets highlight the sensitivity of the users about green spaces in the district. In

ICONARP – Volume 9, Issue 2 / Published: 21.12.2021

Table 3 there are sample tweets of the research which are fed to the model.

Table 3. Sample of tweets by categories

Positive	Botanic garden in Ataşehir…
	The green area is wonderful.
	We dream of botanical garden.
	The best park-garden I have ever seen.
Neutral	We wish green space and parking area.
	They can tear it down and make a great playground.
	A bright future for tomorrow ?? what should you leave to future generations. How about a curse or a prayer? Desert land or green area/forest.
Negative	We want green space, car parking and swimming pool, don't want high-rise buildings that cause traffic congestion.
	Green area is not visible.
	Illegal parking, the lack of green space, the store products on the sidewalk and uniformity problem of colour and/or form in concrete work creates a terrible image.
	The green area was occupied.

User opinions and demands were found to be compatible with the current park and green spaces situation of the district. The percentage of negative comments points to improper situations in planning and design. Since the size of the green area of the district is not at the level of the targeted standards, user dissatisfaction has increased. Users chose to share their expectations and criticisms for public and green spaces through Twitter application with other residents, local administrators and politicians. This dissatisfaction was revealed by user sentiment analysis as a result of data mining. Green space design projects and samples which were appreciated in the district were shared through social media and similar practices were requested.

DISCUSSION AND CONCLUSION

In this study, the total area of green spaces of the district is detected as insufficient against the targeted standards by the map processing method developed for this research. By user sentiment analysis on social media application Twitter, it is concluded that the user thoughts and demands are compatible with this result. The users mostly (72%) have complaints and dissatisfaction about green spaces. The developed applications help to support and check both the text and image data at the same time. This study enables to mine demands and thoughts of the users about green spaces of the district and evaluate this data by matching it with the actual condition of green areas. The applications and the methodology will help to promote citizen participation in taking decisions and design issues about urban planning and architecture. City planners, designers, and politicians can benefit from the information obtained from users of the district regarding their needs and the

DOI: 10.15320/ICONARP.2021.181

inadequate green areas. In this study, a limited data set for processing and sample location is taken into consideration to evaluate. In future studies, social media can be used widely and locations can vary for other related research about urban issues such as water supplies or built environment.

Improving the spatial and urban quality of life will be realized through accurate and well planned designs. A sufficient planning infrastructure is required for the healthy use of all spaces, structures and environments by citizens. Successful designs should be able to meet user needs and demands within sustainability and accessibility requirements. The direct impact of green spaces on urban life and public health and their positive contribution to environmental quality will increase the daily living standards of users. In order to determine the characteristics of these items and to analyse them spatially, it is important to determine user satisfaction. Based on the analysis of different living spaces, it will be possible to reach planning and design ideas throughout the city. Evaluating the user demands before the design stage will be a guide for healthy planning. Reaching user opinions and suggestions about existing designs promotes the collaboration for city development. Social and political areas will benefit from the positive contribution of user participation in planning and design stages. By the result of this research; it has been pointed out that mining data through social media can be used as an effective tool for planning decisions in the field of design. User data can be obtained through current and advanced information technologies. Social media has become an important medium as it provides users cross-sections from their daily lives and mediates the rapid dissemination of information. As in this case study, user sentiment analysis helps making sense of thousands of tweets written on social media. It is also observed that image processing methods can be used to find vegetation on maps and to automate significant detection. Processing and interpreting this data in a proper way will bring new dimensions for design issues. The developing trend of social media in social life enables these studies to be carried out more widely and in detail.

Researchers from different disciplines will be able to contribute to design techniques by using data mining methods in architecture, urban modelling and planning. Analysing and interpreting design products, urban experiences and utilizing these experiences, thoughts, opinions and demands in new products and spatial designs will increase project success. Finding out how the products and designs presented to the users are perceived by them, how practical and real-life problems arise will generate useful feedback for all actors. Reaching the problems, needs and demands of the dwellers will become easier and more practical with these scientific research techniques. Perception and interpretation of the entire built environment and its products by users will be valuable for designers and politicians. It will be aimed to develop different support tools in terms of spatial planning and design. As in the case study, users have dissatisfaction with parks and green areas planning and design within the district. A participatory planning and design will bring positive results in design processes by taking into account the spatial suggestions and demands of the users. It will be possible to develop policies for architectural and urban spaces and to make long-term planning decisions by using user demands and feedback. The researchers can include these outcomes in decision-making processes of planning. The municipalities and designers can benefit from extraction process of user opinions.

CONFLICT OF INTEREST

No conflict of interest was declared by the authors.

FINANCIAL DISCLOSURE

The authors declared that this study has received no financial support.

ETHICS COMMITTEE APPROVAL

Ethics committee approval was not required for this article.

LEGAL PUBLIC/PRIVATE PERMISSIONS

In this research, the necessary permissions were obtained from the relevant institutions and organizations during the study.

REFERENCES

Allahyari, M., Pouriyeh, S., Assefi, M., Safaei, S., Trippe, E. D., Gutierrez, J. B., & Kochut, K. (2017). A Brief Survey of Text Mining: Classification, Clustering and Extraction Techniques. *ArXiv:1707.02919 [Cs]*. http://arxiv.org/abs/1707.02919

Almeer, M. H. (2012). Vegetation Extraction from Free Google Earth Images of Deserts Using a Robust BPNN Approach in HSV Space. *International Journal of Emerging Technology and Advanced Engineering*, *2*(5), 1–8.

Amasyali, M. F., Tasköprü, H., & Çaliskan, K. (2018). Words, Meanings, Characters in Sentiment Analysis. 2018 Innovations in Intelligent Systems and Applications Conference (ASYU), 1–6. https://doi.org/10.1109/ASYU.2018.8554037

Ataşehir Municipality. (2020, September 18). *Tarihçe*. https://www.atasehir.bel.tr/tarihce

Baykara, M., & Gürtürk, U. (2017). Classification of social media shares using sentiment analysis. 2017 International Conference on Computer Science and Engineering (UBMK), 911–916. https://doi.org/10.1109/UBMK.2017.8093536

Brabham, D. C. (2009). Crowdsourcing the Public Participation Process for Planning Projects. *Planning Theory*, *8*(3), 242–262. https://doi.org/10.1177/1473095209104824

Chang, M. C., Bus, P., & Schmitt, G. (2017). Feature Extraction and K-means Clustering Approach to Explore Important Features of Urban

DOI: 10.15320/ICONARP.2021.181

Identity. 16th IEEE International Conference on Machine Learning and Applications (ICMLA), 18-21 December 2017, Pp. 1139-1144, Cancun, Mexico. https://doi.org/10.1109/ICMLA.2017.00015

Chang, M. C., Bus, P., Tartar, A., Chirkin, A. M., & Schmitt, G. (2018). Big-Data Informed Citizen Participatory Urban Identity Design. Proceedings of the 36th Education and Research in Computer Aided Architectural Design in Europe (ECAADe 2018), 19-21 September 2018, Pp. 669-678, Lodz, Poland.

Chatterjee, A., & Perrizo, W. (2015). Classifying Stocks using P-Trees and Investor Sentiment. Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 25-28 August 2015, Pp. 1362-1367, Paris, France. https://doi.org/10.1145/2808797.2808845

Chen, N. C., Nagakura, T., & Larson, K. (2016). Social Media as Complementary Tool to Evaluate Cities—Data Mining Innovation Districts in Boston. A. Herneoja, T. Österlund. and P. Markkanen (Eds.), Complexity & Simplicity - Proceedings of the 34th ECAADe Conference - Volume 2, University of Oulu, 22-26 August 2016, Pp. 447-456, Oulu, Finland.

http://papers.cumincad.org/cgi-bin/works/paper/ecaade2016_096

Chen, N. C., Zhang, Y., Stephens, M., Nagakura, T., & Larson, K. (2017). Urban Data Mining with Natural Language Processing: Social Media as Complementary Tool for Urban Decision Making. G. Çagdas, M. Özkar, L. F. Gül and E. Gürer (Eds.) Future Trajectories of Computation in Design [17th International Conference, CAAD Futures 2017, Proceedings / ISBN 978-975-561-482-3], 12-14 July 2017, Pp. 101-109, Istanbul, Turkey. http://papers.cumincad.org/cgi-bin/works/paper/cf2017_101

Doran, D., Severin, K., Gokhale, S., & Dagnino, A. (2015). Social Media Enabled Human Sensing for Smart Cities. *AI Communications*, *29*(1), 57–75. https://doi.org/10.3233/AIC-150683

Elragal, A. (2014). ERP and Big Data: The Inept Couple. *Procedia Technology*, *16*, 242–249. https://doi.org/10.1016/j.protcy.2014.10.089

Google. (2021, August 6). *FAQ*. https://earthengine.google.com/faq/

Han, J., Kamber, M., & Pei, J. (2012). Introduction. In J. Han, M. Kamber, & J. Pei (Eds.), *Data Mining: Concepts and Techniques (Third Edition)* (pp. 1–38). Morgan Kaufmann. https://doi.org/10.1016/B978-0-12-381479-1.00001-0

Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Ullah Khan, S. (2015). The rise of “big data” on cloud computing: Review and open research issues. *Information Systems*, *47*, 98–115. https://doi.org/10.1016/j.is.2014.07.006

Hasib, K., Habib, Md. A., Towhid, N. A., & Showrov, Md. I. H. (2021). A Novel Deep Learning based Sentiment Analysis of Twitter Data for US Airline Service. International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD), 27 February 2021, Pp. 450-455, Dhaka, Bangladesh. https://doi.org/10.1109/ICICT4SD50815.2021.9396879

Hegadi, R., & Sangolli, R. (2011). Segmentation of Google Map Images Based on Color Features. Second International Conference on Communication, Computation, Management & Nanotechnology (ICN-2011), 23-25 September 2011, Bhalki, India.

Joulin, A., Grave, E., Bojanowski, P., & Mikolov, T. (2016). Bag of Tricks for Efficient Text Classification. *ArXiv:1607.01759 [Cs]*. http://arxiv.org/abs/1607.01759

Kaiser, P., Wegner, J. D., Lucchi, A., Jaggi, M., Hofmann, T., & Schindler, K. (2017). Learning Aerial Image Segmentation from Online Maps. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(11), 6054–6068. https://doi.org/10.1109/TGRS.2017.2719738

Khan, A., Atique, M., & Thakare, V. M. (2015). Combining Lexicon-based and Learning-based Methods for Twitter Sentiment Analysis. *International Journal of Electronics, Communication and Soft Computing Science & Engineering (IJECSCSE)*, 89–91.

Khan, N., Yaqoob, I., Hashem, I. A. T., Inayat, Z., Mahmoud Ali, W. K., Alam, M., Shiraz, M., & Gani, A. (2014). Big Data: Survey, Technologies, Opportunities, and Challenges. *The Scientific World Journal*, *2014*, e712826. https://doi.org/10.1155/2014/712826

Kitchin, R. (2014). The real-time city? Big data and smart urbanism. *GeoJournal*, *79*(1), 1–14. https://doi.org/10.1007/s10708-013-9516-8

Kobayashi, Y., Munezero, M., & Mozgovoy, M. (2016). Analysis of Emotions in Real-time Twitter Streams. *Informatica*, *40*(4), Article 4. https://www.informatica.si/index.php/informatica/article/view/1462

Kotzias, D., Lappas, T., & Gunopulos, D. (2016). Home is where your friends are: Utilizing the social graph to locate twitter users in a city. *Information Systems*, *57*, 77–87. https://doi.org/10.1016/j.is.2015.10.011

Lessmann, S., Baesens, B., Seow, H.-V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research*, *247*(1), 124–136.

Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., & Zhang, W. (2015). Assessing street-level urban greenery using Google Street View and a modified green view index. *Urban Forestry & Urban Greening*, *14*(3), 675–685. https://doi.org/10.1016/j.ufug.2015.06.006

Ministry of Environment and Urbanization. (2019, August 6). *Mekansal Planlar Yapım Yönetmeliği*. https://mpgm.csb.gov.tr/mekansal-planlar-yapim-yonetmeligi-i-4583

Mueller, J., Lu, H., Chirkin, A., Klein, B., & Schmitt, G. (2018). Citizen Design Science: A strategy for crowd-creative urban design. *Cities*, *72*, 181–188. https://doi.org/10.1016/j.cities.2017.08.018

Münster, S., Georgi, C., Heijne, K., Klamert, K., Rainer Noennig, J., Pump, M., Stelzle, B., & van der Meer, H. (2017). How to involve inhabitants in urban design planning by using digital tools? An overview on a state of the art, key challenges and promising approaches. *Procedia Computer Science*, *112*, 2391–2405. https://doi.org/10.1016/j.procs.2017.08.102

DOI: 10.15320/ICONARP.2021.181

Nguyen, H. L., Woon, Y. K., & Ng, W. K. (2014). A Survey on Data Stream Clustering and Classification. *Knowledge and Information Systems*, *45*. https://doi.org/10.1007/s10115-014-0808-1

OpenStreetMap. (2021, June 1). Copyright. https://www.openstreetmap.org/copyright/en

Pan, Y., Tian, Y., Liu, X., Gu, D., & Hua, G. (2016). Urban Big Data and the Development of City Intelligence. *Engineering*, *2*(2), 171–178. https://doi.org/10.1016/J.ENG.2016.02.003

Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*, *2*(1–2), 1–135. https://doi.org/10.1561/1500000011

Psyllidis, A., Bozzon, A., Bocconi, S., & Titos Bolivar, C. (2015). A Platform for Urban Analytics and Semantic Data Integration in City Planning. In G. Celani, D. M. Sperling, & J. M. S. Franco (Eds.), *Computer-Aided Architectural Design Futures. The Next City—New Technologies and the Future of the Built Environment* (Vol. 527, pp. 21–36). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-47386-3_2

Rao, T., Mitra, P., Bhatt, R., & Goswami, A. (2018). The big data system, components, tools, and technologies: A survey. *Knowledge and Information Systems*. https://doi.org/10.1007/s10115-018-1248-0

Resch, B., Summa, A., Zeile, P., & Strube, M. (2016). Citizen-Centric Urban Planning through Extracting Emotion Information from Twitter in an Interdisciplinary Space-Time-Linguistics Algorithm. *Urban Planning*, *1*(2), 114–127. https://doi.org/10.17645/up.v1i2.617

Salesses, P., Schechtner, K., & Hidalgo, C. A. (2013). The Collaborative Image of The City: Mapping the Inequality of Urban Perception. *PLoS ONE*, *8*(7), e68400. https://doi.org/10.1371/journal.pone.0068400

Salur, M. U., & Aydin, I. (2020). A Novel Hybrid Deep Learning Model for Sentiment Classification. *IEEE Access*, *8*, 58080–58093. https://doi.org/10.1109/ACCESS.2020.2982538

Santos, C. d, & Gatti, M. (2014). Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts. Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers, 69–78. https://aclanthology.org/C14-1008

Schmidt, A., & Wiegand, M. (2017, April). A Survey on Hate Speech Detection using Natural Language Processing. Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media, 3-7 April 2017, Pp. 1-10, Valencia, Spain. https://doi.org/10.18653/v1/W17-1101

Seltzer, E., & Mahmoudi, D. (2013). Citizen Participation, Open Innovation, and Crowdsourcing: Challenges and Opportunities for Planning. *Journal of Planning Literature*, *28*(1), 3–18. https://doi.org/10.1177/0885412212469112

Tasse, D., & Hong, J. I. (2014). Using Social Media Data to Understand Cities. *Carnegie Mellon University, Journal Contribution*. https://doi.org/10.1184/R1/6470645.v1

Thakuriah, P., Tilahun, N., & Zellner, M. (2017). *Big data and urban Informatics: Innovations and challenges to urban planning and knowledge*

ICONARP – Volume 9, Issue 2 / Published: 21.12.2021

discovery (P. Thakuriah, N. Tilahun, & M. Zellner, Eds.; pp. 11–45). Springer. http://www.springer.com/gb/book/9783319409009

TUIK. (2019, August 8). *Calculation of Address Based Population Registration System Results and Publication in Central Distribution System*. https://biruni.tuik.gov.tr/medas/?kn=95&locale=tr.

Valls, F., Redondo, E., Fonseca, D., Kompen, R. T., Villagrasa, S., & Martí, N. (2018). Urban data and urban design: A data mining approach to architecture education. *Telematics Informatics*, *35*, 1039–1052. https://doi.org/10.1016/J.TELE.2017.09.015

Wong, T.-T. (2015). Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition*, *48*(9), 2839–2846. https://doi.org/10.1016/j.patcog.2015.03.009

World Cities Culture Forum. (2019, August 6). *% of public green space (parks and gardens)*. http://www.worldcitiescultureforum.com/data/of-public-green-space-parks-and-gardens

Yang, C., Huang, Q., Li, Z., Liu, K., & Hu, F. (2017). Big Data and cloud computing: Innovation opportunities and challenges. *International Journal of Digital Earth*, *10*(1), 13–53. https://doi.org/10.1080/17538947.2016.1239771

Younas, M. (2019). Research challenges of big data. *Service Oriented Computing and Applications*, *13*(2), 105–107. https://doi.org/10.1007/s11761-019-00265-x

Zikopoulos, P. C., Eaton, C., deRoos, D., Deutsch, T., & Lapis, G. (2012). What is Big Data? Hint: You're a Part of It Every Day. In *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data* (pp. 3–13). McGraw-Hill.

Resume

Hülya Soydaş Çakır has been Assistant Prof. in the Faculty of Engineering and Architecture at Fenerbahçe University since 2018. She received her bachelor's and master's degree in Architecture from İstanbul Technical University. She had Ph.D. in Informatics from Marmara University. She has architectural experience in various projects, publications and researches on Digital Learning and Design Environments, Computer Aided Design, Education, Universal Design, Information Technologies and Design Disciplines.

Vecdi Emre Levent has been Assistant Prof. in the Faculty of Engineering and Architecture at Fenerbahçe University since 2019. He received his bachelor's degree from Arel University, his master's degree from Yıldız Technical University and his Ph.D. from Özyeğin University in Computer Engineering. His experience includes computer arithmetic and architecture, VLSI/FPGA design and automation, embedded systems, machine vision and image processing. He provides consultancy for various defence industry companies.

DOI: 10.15320/ICONARP.2021.181