Research Article

Content Analysis on Homelessness Issues at Airports by News Media Mining

Subasish Das[1] ⓘD, **Javad J. C. Aman[2]** ⓘD, and **M. Ashifur Rahman[3]** ⓘD

Abstract

The number of homeless people at airports has increased in recent years. As airports are safe, transit-accessible, convenient, and climate-controlled facilities with food and amenities, these places are attractive to homeless people who need a safe and secure place to stay. The main struggle of airports in this regard is maintaining a balance between customers, who are mostly the traveling public, and dealing with homeless people delicately. Moreover, because of their poverty and insufficient or no access to healthcare, these people suffer from physical and mental issues. With the COVID-19 pandemic, this problem became more critical. Many news media outlets started to report on homelessness at airports. News-framing impacts have some contribution in the context of this issue. However, the impact of news coverage on "airport and homelessness" has not yet been studied. News-framing effects have been identified in the context of tourist destinations. Although many studies have explored homelessness and transit, this issue at airports has not been well studied. This study provides a brief overview of the issue of homelessness in the transportation domain, including transit and aviation. Additionally, this study collected news articles related to "airport and homelessness" (71 articles) both during the COVID-19 pandemic (March 1, 2020–July 21, 2021) and before the pandemic (before March 1, 2020). These news articles contain around 50,000 words. As the data is unsupervised in nature, a text network analysis was performed to determine the latent information from these textual contents. The findings of this study can shed some light on this scientifically unexplored but widely discussed issue.

As the number of homeless persons within the U.S.A. is growing, issues surrounding health, patronage, safety, and costs of the transportation industry are appearing. The areas that are most strongly affected include transit, air, and rail facilities. For example, Schwartz (*1*) examined how the characteristics of homeless people have changed. This study looked at how the transportation industry has been involved in both increasing the problem and addressing it with various programs. As the transportation industry is growing, it is anticipated that the problem of the relationship between transportation facilities and homeless people will expand. It is expected the transportation industry will take a more active but costly role as it is both a single entity but also operates in tandem with other service agencies.

As airports are safe, transit-accessible, convenient, and climate-controlled facilities with food and amenities, they are attractive places to homeless people who need a safe and secure place to stay. The main struggle of the airports in this regard is to maintain a balance between customers, who are mostly the traveling public, and dealing with homeless people delicately. Moreover, because of their poverty and insufficient or no access to healthcare, homeless people often suffer from physical and mental issues. These mental health conditions could also arise from substance abuse and may make them a threat to themselves or others. Consequently, the long-term presence of homeless people could lead to a rise in criminal activities such as assault, theft, and vandalism. This

[1]Texas A&M Transportation Institute, Bryan, TX
[2]Sharif University of Technology, Tehran, Iran
[3]University of Louisiana at Lafayette, Lafayette, LA

Corresponding Author:
Subasish Das, s-das@tti.tamu.edu

could further extend to jeopardizing the safety of travelers in the airport. With the COVID-19 pandemic, this problem has been more critical. A review of the literature shows that "homelessness and transportation" have been well investigated to some extent, but less focus is given to "airport and homelessness," although this issue has been exacerbated by the COVID-19 pandemic. On the other hand, multiple news agencies have found this phenomenon important enough to reflect on and investigate. The literature has also highlighted the importance of news articles as an insightful source of information in the first stages of investigating a social phenomenon. Therefore, the goal in this study is to conduct a text mining analysis to shed light on different dimensions of the homelessness issue at airports.

News media (mostly local and some national news outlets) started to cover "airport and homelessness" more extensively during the COVID-19 pandemic. Online news articles are regarded as one of the common sources for reflecting current events and, moreover, news agencies mostly try to reflect the topics of interest to their society. News coverage can not only deeply investigate social phenomena and offer legitimate solutions but also sometimes form communities' points of view on the events triggering the reporting (*2*). Therefore, analyzing what news media cover and suggest seems a critical first step in investigating social phenomena. In the world of social media and the plethora of online media, the delivery of news seems to be associated with different discourses. However, there is little investigation on how media coverage on this issue is impactful in understanding and solving this issue.

An overwhelming number of news articles can be found on the internet, shared by different news agencies. Moreover, textual datasets often lack metadata such as topic, semantic structure, and labels, and frequently include noise. This makes news articles one of the most difficult datasets for analysis and investigation. Researchers usually use text mining approaches to explore topical patterns within news articles automatically. This study performed a thorough literature review and news media text mining by collecting relevant news articles on this issue. Relevance was assessed by the news articles answering two major research questions. First, is there any difference in news coverage between the COVID-19 pandemic and the pre-pandemic period? Second, what are the relevant topics in the news coverage on "airport and homelessness"? The study answered these questions by performing text network analysis and collocation (i.e., group of words or phrases) analysis with inclusion of relevant news media articles.

Literature Review

There has not been much research on the magnitude and operative consequences of homelessness for the aviation sector. Additionally, no literature has confronted the topic on what would be considered the "best practices" for resolving this issue. It is important to provide a literature review on this issue by considering overall transportation sectors. The findings of this literature review can provide contexts related to airports and the aviation sector.

Transportation and Homelessness

Although the bulk of research on homelessness has been conducted on various issues such as mental health and economic conditions, a portion has focused on exploring transportation-related issues and homelessness. In the U.S.A., departments of transportation are heavily involved in dealing with the consequences of homelessness. Bassett et al. (*3*) investigated responses to homeless encampments that were located on rights-of-way owned by Oregon Department of Transportation (DOT). The prevalence of the issue was examined, information that could be useful in the development of a best practices guide was gathered from the data, and the responses of DOT staff were recorded. The research utilized a single mixed-methods, electronic survey of practitioners, an in-depth case study, and follow-up interviews.

It is difficult to find guidance for the policy-making process with regard to homeless individuals' travel needs because of the lack of research on the intersection of transportation and homelessness, along with the lack of understanding of travel behavior and mobility needs of homeless individuals. Hui and Habib (*4*) designed and conducted an interview across eight shelters and drop-in agencies in Toronto, Canada with 159 homeless individuals. Stated adaptation experiments were included to determine how homeless individuals' activity participation processes were affected by transportation. Using econometric modeling techniques, decision choice models were estimated to demonstrate the influence of various variables on the decision to reject or accept employment opportunities (such as age, income, and duration of homelessness). The choice to travel for social activity was influenced by a person's typical mode selection and daily trip frequency.

Boyle (*5*) discussed beneficial practices, conclusions, and methods on relations between homeless people and the transportation industry in a literature review of records from the 1980s to the present day. The literature review additionally analyzes procedures and guidelines of public libraries with regard to homeless individuals, since public libraries, which are safe places for the homeless, provide their services to all public members despite their situation. A network-based survey of transport organizations in North America was issued to record current problems and practices in relations between homeless

individuals and transportation organizations, which also included examinations of factors that caused several strategies to be either successful or unsuccessful. The research also included six comprehensive cases that describe successful and inventive methods. It was shown that homeless individuals are considered a problem for transport organizations whatever the size of the organization, even though larger organizations usually categorized homelessness as a major problem. It was also discovered that successful procedures should focus on behavior instead of targeting communities or individuals, and constant reinforcement and codes of conduct refine the expectations of these transit organizations. Besides that, collaborations between agencies were shown to be effective, and that enforcement is required but not sufficient for success. These collaborations improved customer safety and perceptions, increased sensitivity toward the people and the topics involved, and provided aid to those who needed it.

Service Accessibility and Homelessness

The connections between transportation and societal exclusion illustrate the importance of public transit as a mode of transportation for people who are unable to access a private mode. In this study, people most at risk of being at the furthest end of societal exclusion and users of public transit are low-income and homeless people, as most of the time their only form of transportation is by public transit. Despite this identification in the study, this community of at-risk people are rarely the focus in the application and establishment of transportation guidelines. Hui and Habib (6) mainly focused on the problems of the at-risk group facing transportation-related exclusion as well as issues with public transport service accessibility in the city of Toronto. The foundation of the study was the rate of public transportation services being the main factor defining transportation-related social exclusion in a survey sample that was issued out to a group of low-income and homeless people in Toronto. The researchers also investigated current transport guidelines in cities and voluntary and community transport service workers, which asserted the need for policy changes to improve the inclusion of the at-risk group in transportation development, as well as discounted prices for transit fares for certain groups, increased policy integration among the various levels of government, and better transit accessibility for low-income districts and neighborhoods.

Service accessibility intended to assist homeless, socially underprivileged, and low-income individuals can often be problematic, even more so for people who often do not have access to private modes of transportation. This makes transport accessibility services vital to people who depend on these transport services for daily life. To measure transport accessibility, Lamers et al. (7) calculated the time transit services at a city in Toronto took to arrive at several homeless shelters and each census region. Using four attributes from census data, an indicator of social disadvantage was used to score each census region, and it was discovered that because of the more frequent transport supply and the high concentration of services, the old City of Toronto had high access levels. However, in the outer areas of the city, some places faced the issue of having concerning levels of access to community services with regard to transportation and accessibility. The travel time matrices that were provided from this study can be used as guidelines for future service developments with regard to procedural and policy analysis.

Grimes and Smirnova (8) investigated the effects of bicycle ownership on homelessness-related mobility, health, self-esteem, social capital, and employment, with a particular focus on how the bicycle was utilized by the participant and how the bicycle affected perceived physical, emotional, and financial independence. As these factors may contribute to an increase in self-esteem, this was the focus of the study. Sixteen men experiencing homelessness who had graduated from a local homeless outreach agency's earn-a-bike program were given in-depth, open-ended interviews that were recorded, transcribed, and analyzed using Dedoose software. Multiple positive effects, such as maintaining employment, improving health and access to health care services, strengthening social capital, and boosting self-esteem, were found to result from the men who were experiencing homelessness acquiring bicycles and valuing the increased independence bicycles afforded them. This study provided insights into a cost-effective, sustainable intervention for the improvement of men experiencing homelessness' transportation, health, employment opportunities, social capital, and self-esteem. This study also highlighted obstacles to bicycle ownership, some of which can be mitigated through the provision of locks to secure the bicycles.

A study that validated a scale which measures transportation accessibility among homeless individuals (who are some of the heaviest users of public transport in the U.S.A.) in low-density urban communities was covered by Cronley et al. (9). In spring 2018 and winter 2019, surveys that included a nine-item, four-point Likert scale were administered (N = 164; 38% African American). Higher scores indicated greater difficulties in access to transportation. A confirmatory factor analysis (CFA) and an exploratory factor analysis were conducted. The mean differences in the transportation disadvantage (TD) scale between a community with mass transit and a community without mass transit were tested to evaluate

criterion validity. A two-factor solution arose from the CFA: six items addressing costs, accessibility, and the ability to locate public transport were included in factor one, whereas two items addressing physical accessibility were included in factor two. Because a poorly fitting model was produced by its inclusion, the item that measured perceptions of safety for traveling was not included. The community without mass transit had respondents who scored higher. Adequate psychometric properties for application with individuals experiencing homelessness was demonstrated by the TD scale. Females were statistically shown by post-hoc analysis to have more difficulty with regard to safety than males.

Mobility and Homelessness

The mobility of the homeless is often constrained by certain geographic and social factors. The concentration of the homeless in public spaces in downtown areas that can be used for sleeping and networking, and neighborhoods that have high levels of poverty may be geographically identified as "spaces of containment." For homeless people to be able to move between stigmatized and non-stigmatized places, mobility is essential; subsequently, public transit is important for homeless individuals to be able to access services, education, shelter, affordable housing, and employment. The mobility patterns and public transit use of homeless people located in Long Beach, CA was studied by Jocoy and Del Casino (10). The authors also examined how public transit affects the ability of homeless persons to navigate between places.

Jocoy and Del Casino (11) observed the day-to-day mobility of homeless adults, again in Long Beach. This was done to determine how the daily travel of homeless people compares with their imagined characteristics and with the national household travel patterns of the United States. Through the use of focus groups and structured interviews, it was determined that the mobility of the homeless is highly spatially constrained and structured by sociocultural relations. These sociocultural relations consist of stigmatization, individual accountability, and economic output.

COVID-19 and Homelessness

The homelessness epidemic is a serious issue as it usually exceeds the capabilities of the current safety net—even more so since the COVID-19 pandemic has amplified this issue. Many homeless people resort to using transport vehicles, stations, and transit stops as a means of temporary shelter. Besides transit being used as shelter, it can also be used to get to certain places like work facilities, other shelters, and public service facilities. Loukaitou-Sideris et al. (12) analyzed the relations between transport, homelessness, and the COVID-19 pandemic, showcasing the outcomes of a survey of 115 transport operatives on their perceptions of how homelessness affects their organizations. Based on the survey, homelessness was shown to be present among several transport organizations, but was mainly focused on larger organizations and key hotspots, as well as the pandemic worsening the capacity of the transport agencies. To improve the results and transit experience for both homeless riders and the general public, in fulfilling their social service functions, transportation organizations should focus on the wellbeing and mobility of homeless users.

Potential Solutions

Loukaitou-Sideris et al. (13) investigated the magnitude of the effect of homelessness on transport and how transport organizations are resolving this issue. Part I describes the degree of homelessness on transport in various locations by integrating prior research and count data. Transport areas are also used as means of housing for a large yet varying amount of unhoused people, who are more often structurally underprivileged and chronically homeless than other individuals who find shelter in other places. Part II presents comprehensive case studies of strategies used by transport organizations across the U.S.A., which are as follows: service hubs, reduced fare prices, mobile outreach, and transit to shelters. The magnitude, application, effect, complications, and lessons learned of the various strategies were also presented. Gathering data more methodically, limiting the routine interference by law enforcement officers with the homeless, encouraging external collaborations, public education, and training transportation workers were shown to be beneficial.

Several factors including alcoholism, drug addiction, unemployment, fewer services for those who use mental health resources, and excessive housing costs, add to the homelessness epidemic in California and across the U.S.A. Faced with these factors, to seek safety, homeless individuals turn to finding any means of cover or shelter that they can and have even resorted to riding buses all night just to keep themselves from the cold and from danger. This is especially seen with homeless individuals who are older, medically disadvantaged, and those who do not want to be subject to violence or other disadvantages that they may face in the streets. In response, angry residents and patrons of the transit systems have elicited several strict policies and anti-camping regulations against homeless individuals, especially in more urban areas. The Orange County Transportation Authority (OCTA) maintains and supplies the Orange County Sheriff's Department with transport police services and law enforcement as stated by the terms of a $4.2 million

contract. Rudy and Delgado (14) discussed an inventive partnership between the Orange County Health Care Agency, the OCTA Transit Police Services, and the Assembly Bill (AB) 2034 Homeless Outreach Program.

Carlson et al. (15) discussed how to fulfill requirements of the McKinney-Veneto Homeless Education Assistance Act, in which the State of Washington provided funds to initial homeless student transit programs from the years 2004 to 2006. The study confronted four questions on these initial efforts to address homelessness among students: what forms of transit were utilized, how much did they cost, which forms were preferred, and did attending the original school influence the students' academic performance? To accomplish this, cost data from eight educational service districts and on readership were accessed, and interviews between homeless students, transit coordinators, parents, and homeless associations were conducted. The school districts used a wide variety of transportation systems to mobilize students and discovered that homeless student transit was often costly and that attending the original school a student had primarily attended was related to better Washington Assessment of Student Learning (WASL) scores.

Bassett and Tremoulet (16) introduced a case study of the transfer of homeless housing units from the Baldock Rest Area near Wilsonville to the urban city of Portland, OR in 2010 to 2011. This report explored the method used in the transfer, explained the procedures and main factors used to resolve this issue, and evaluated the results for public organizations and the homeless individuals who were transferred by gathering information from archive materials, two national surveys of DOT/rest area managers, and crucial informant interviews. What made this relocation successful and was critical to this process was the interagency partnerships and the application of both restrictions (push) and incentives (pulls). Overall, DOT maintenance managers, supervisors, and engineers should actively implement these procedures and methods that humanely and efficiently tackle homelessness on DOT property.

To the best of the authors' knowledge, there is no previous work on the news media discourse using text mining on "airport and homelessness." For that reason, the aim of this study is to unearth insights on this topic through the text mining of news media coverage on the issue.

News Coverage and Text Mining

News is "the communication of selected information on current events," with the selection based on "newsworthiness" or "what the public finds interesting" (17). News is one of the most reliable sources for understanding current events. The news trends can even help researchers to investigate various opinions about recent events. However, the information era has offered an overwhelming number of textual documents available on the internet. Moreover, textual datasets often lack metadata such as topic, semantic structure, and labels (18), and frequently include noise. These makes news articles one of the most difficult datasets for analysis and investigation. Researchers usually use text mining approaches to explore topical patterns automatically within the news articles. Text mining or knowledge investigation is mostly related to the automatic exploration of latent patterns from unstructured textual datasets (19). Literature has shown a broad range of applications for text mining; common applications include sentiment analysis (20), forecasting petitions (21), market prediction (22) and user opinion mining (23). News articles are content-rich resources for exploring qualitative information using descriptive text analysis. However, not many studies, specifically in the field of transportation, have utilized this approach to investigate news articles. This study applied this innovative approach to determine key insights on the topic of airport and homelessness.

Methodology

Data Collection

As news media publish many stories on "airport and homelessness," it is important to know the information provided in these news stories. As the news stories differed during the COVID-19 and pre-COVID periods, the collected news articles were divided into two groups based on the time periods: COVID-19 (March 1, 2020–July 21, 2021) and pre-COVID (before March 1, 2020). This study used Google News Alert by using keywords such as "airport homeless" and "airport homelessness" to collect the relevant URLs (URL, or Uniform Resource Locator, indicates the address of a webpage). After web scraping using the popular Python library Beautiful Soup (https://www.crummy.com/software/BeautifulSoup/bs4/doc/), 71 articles were finally selected for analysis. The selected news articles are shown in Table 1. Of the 71 articles (as these news articles are easily searchable using the article title, urls are not provided in the table), around 70% were published during the COVID-19 pandemic and 30% were published in the pre-COVID period. The pre-COVID period extends to 1995, so the shorter duration of the COVID-19 pandemic period shows more news articles compared with the pre-COVID period. All the news articles were investigated manually to check if their contents were related to the issue of homelessness and airports. Within the list of selected articles, there are a few articles that are not completely dedicated to the mentioned topic. These articles, however, shed light on a part of the story of homelessness and airports, and can potentially assist us to investigate the topic holistically.

Table 1 News Headlines on "Airport and Homelessness"

Date	News headline or title
News articles during the COVID-19 pandemic	
3/5/2020	DFW Airport Says Homeless No Longer Welcome
3/29/2020	Stranded at LAX Since Christmas Eve. Homeless. Then the Coronavirus Hit
4/3/2020	Coronavirus: Heathrow Airport Sheltering "200 Homeless People"
4/7/2020	Crisis at the Airport: Homeless Could Cause Coronavirus Hot Spot at Hartsfield-Jackson
4/7/2020	Atlanta's Homeless Seek Shelter at Airport; Officials Say "Isolation Facility" Close to Opening
4/17/2020	Plan Proposed to Remove Homeless From Atlanta Airport
5/5/2020	Councilman: Allowing Homeless to Hang out at Atlanta Airport Posed a COVID-19 Health Risk
5/5/2020	About 100 Homeless People Staying at Philadelphia's Desolate Airport
5/6/2020	About 100 People Experiencing Homelessness Staying at Philadelphia's Desolate Airport
5/7/2020	Amid Pandemic, City Plan Directs Homeless Sleeping at Airport to Supportive Services
5/8/2020	Homeless Living at Airport—and One Man Found in a Parked Plane—Underscore the Wreckage of COVID-19
5/9/2020	Conflicting Plans on Helping Homeless Creates Friction
5/21/2020	City Agrees to Test Philly Airport Homeless for COVID-19 After Advocates Threaten Suit
5/26/2020	Philadelphia Officials Relocating Homeless People From Airport Terminal
5/26/2020	City Works to Relocate Homeless From Philly Airport
5/26/2020	Homeless People at Idled Airport Venue to be Tested, Moved
5/28/2020	Homeless People "Relocated" From Philly Airport, Test Negative for COVID-19
7/31/2020	Homeless Couple Upset After O'Hare Implements New Rules Restricting Access
8/13/2020	A 150-Bed Homeless Shelter Near the Fullerton Airport Opens
8/26/2020	Columbia Airport Displays Artwork by Residents of Transitions Homeless Center
12/8/2020	Homeless Man Charged With Breaking and Entering Into Vehicles in West Springfield
12/23/2020	"I'm Not Scared Anymore": New Castle County's Hotel-Turned-Homeless Shelter Gets Underway
1/18/2021	Man Lives at Chicago O'Hare Airport for MONTHS Over "COVID Fears"
1/19/2021	A Man Was "Scared to Go Home due to COVID." So He Spent Three Months Living in Chicago Airport, Prosecutors Say
2/7/2021	Woman Dies at Homeless Camp in the Yard of Salt Lake City Activist
3/1/2021	Atlanta Airport's Approach to Homeless Outlined in Proposed Ordinance
3/6/2021	Caitlyn Jenner Is Running for Governor Because Her Private-Jet-Flying Friends Are Sick of the Homeless
3/8/2021	Photos: An Unflinching Look at Homelessness During the Pandemic
3/26/2021	Louisville Homeless Encampments Given 21-Day Notice to Vacate
4/6/2021	Terminal Dwellers: the Stories of People Living at the Airport
4/26/2021	"Stop Stealing From Us": Victorville Homeless Shelter Asks Public to Help Find Second Stolen Van
4/27/2021	Homeless Staying in Kansas City Hotels Through Temporary Program now Being Moved
5/24/2021	Louisville to Clear Homeless Camps Downtown, Near Airport Before Kentucky Derby
6/6/2021	Facing Removal from Mineta-SJ Airport, Homeless Campers Moving to Land Owned by Apple
6/14/2021	San Jose Business Owner Fed Up After Repeat Burglaries by Homeless Near Airport
6/15/2021	Apple in Talks With San Jose Over Homeless Encampment on its Undeveloped Land
6/22/2021	Over 500 Homeless Could Be Sheltered Near Chico Airport
6/25/2021	Temporary Resting Site Opens Near Chico Airport
6/25/2021	People Without Homes Voice Frustration, Fear About City Shelter Solution
6/28/2021	FAA Raises Concerns Over Sprawling Homeless Encampment Near San Jose Airport
6/28/2021	Darwin Airport Inundated as Travelers Rush to Flee COVID-19 Lockdown Extension
6/29/2021	FAA Demands Cleanup of Airport-Adjacent Homeless Camp, San Jose Asks for Extension
7/1/2021	Federal Agency Demands Closure of Massive San Jose Homeless Camp
7/9/2021	A Stalemate Continues Over the Chico Homeless Crisis
7/12/2021	"He Probably Figured He'd Tough it out": Homeless Man Died in Heat Wave, Minutes From Cooling Center
7/12/2021	"Heat a Different Animal" for Las Vegas Homeless
7/18/2021	People Vacate Homeless Encampments on Market Street Ahead of Sweep
7/18/2021	State of Homelessness	Tehama County Pushes Forward in Homeless Services
7/21/2021	Hundreds Left Homeless in Douala After Demolition of Airport Neighborhood
News Articles before COVID-19	
2/3/1995	Airport Homeless: A Long, Pleasant Layover
7/21/2000	Homeless Women Seeking Refuge at Logan Airport
6/3/2008	Homeless and Living at the Airport
3/16/2009	Airports' Homeless Resist Efforts to Remove Them
8/27/2014	Homeless Woman at San Jose Airport Boards Plane Without a Ticket
12/27/2014	Homeless People Sleeping, Bathing at Orlando International Airport
3/10/2015	Homeless Man Lived at LaGuardia Airport for 20 Years
3/10/2015	NYC's Homeless Turning to LaGuardia Airport for Safe Shelter
3/30/2015	Overnight Airport Lobby Closure Cuts Down on Homeless

(continued)

Table 1 (continued)

Date	News headline or title
12/4/2015	No Room at the Airport
12/27/2015	Homeless People Not Welcome in Airports as Officials Tighten Rules
2/9/2017	LAX Is Scrambling to Relocate Hundreds of Homeless People Living in Manchester Square
4/26/2017	Homeless Campers Putting Down Roots Near Airport Businesses
2/12/2018	Atlanta's Homeless Fill Atrium of World's Busiest Airport At Night
5/19/2018	Airport Will Collect Change for Homeless
8/1/2018	Atlanta Airport Seeks Ok for Contract to Move Homeless From Terminal
1/25/2019	Homeless Young People Are Hiding All Over Central Florida — Including at the Airport
4/22/2019	Homeless, at the Airport
8/21/2019	Housing Insecurity in the Nation's Richest Cities Is far Worse Than Government Statistics Claim
11/2/2019	JetBlue Apologizes After Fort Lauderdale Airport Employee Dresses as Homeless Person
1/7/2020	Airports Shutting Terminals at Night in Effort to Evict Homeless
2/28/2020	Homeless Being Turned Away from DFW Airport

COVID-19 affected people's lives significantly. With travel restrictions and safety protocols, many people started to lose their jobs during the COVID-19 pandemic. As many people could no longer afford to pay the mortgage or rent on their homes during the economic downturn during the COVID-19 pandemic, many were evicted. Seeing no other viable options, some proportion of these homeless people moved to airports to live. The news media, especially the local news media, started to cover this issue frequently. Some news media produced very detailed news stories (maximum word count in a news story = 5,274), and some news stories were very brief (minimum word count in a news story = 49). Figure 1 shows the results of statistical significance tests of the word counts in news titles and news stories for these two periods: COVID-19 pandemic and pre-COVID. R software (https://www.r-project.org) was used for all statistical analysis. The library "*ggstatsplot*" (https://cran.r-project.org/web/packages/ggstatsplot/index.html) was used to perform the significance tests of these differences. Figure 1a shows that the word counts in news titles/headlines during the pre-COVID period have a significantly lower mean count ($\mu = 9.00$) in comparison with the word counts in news headlines ($\mu = 11.02$) during COVID-19. Figure 1b shows that the word counts in news stories during the pre-COVID period had a higher mean count ($\mu = 845.50$) in comparison with the word counts in news stories ($\mu = 585.86$) during COVID-19. However, this difference is not statistically significant. The trends of pre-COVID and COVID-19 news story lengths indicate that the most detailed news on "homelessness at airport" were not published during the COVID-19 period. In many cases, the local news media during COVID-19 covered local news and provided some details, which were often not thoroughly reviewed and vetted through a critical news review process.

Results and Discussion

Text mining is one of the popular data mining methods. It is a branch of natural language processing (NLP) that provides exploratory analysis of unstructured datasets. Text mining can be extended to complicated statistical analysis based on the research question or research needs. Text network analysis, a tool of text mining, can provide important insights on hidden trends of the unstructured text data. This method consists of three major processes:

- Step 1. Text cleaning and token development: The first steps require some required text mining steps to prepare the data for analysis. These steps include text cleaning by using (i) stop word removal and removal of study-specific needs such as removal of redundant punctuation, number, special text or URLs, (ii) token or phrase generation, and (iii) perform stemming (i.e., word root generation; for example, the stem of 'injury' and 'injuries' is 'injur') or lemmatization (i.e., word generation by morphological analysis; for example, the lemma of 'injury' and 'injuries' is 'injury') based on the study needs.
- Step 2. Node and edge generation: This step involves a word search algorithm. The algorithm will first find two words or *n* words sentence by sentence in a corpus (i.e., a document with words and sentences). For the first combination, node with a word and edge between two words or *n* words will be created. The search will continue to acquire such pair or pairs for following sentences. If such pair or pairs exist, the weight on node and edge will crease by a unit weight. With the presence of a new word or new pair, the network will grow and follow the steps of weighting throughout the corpus.
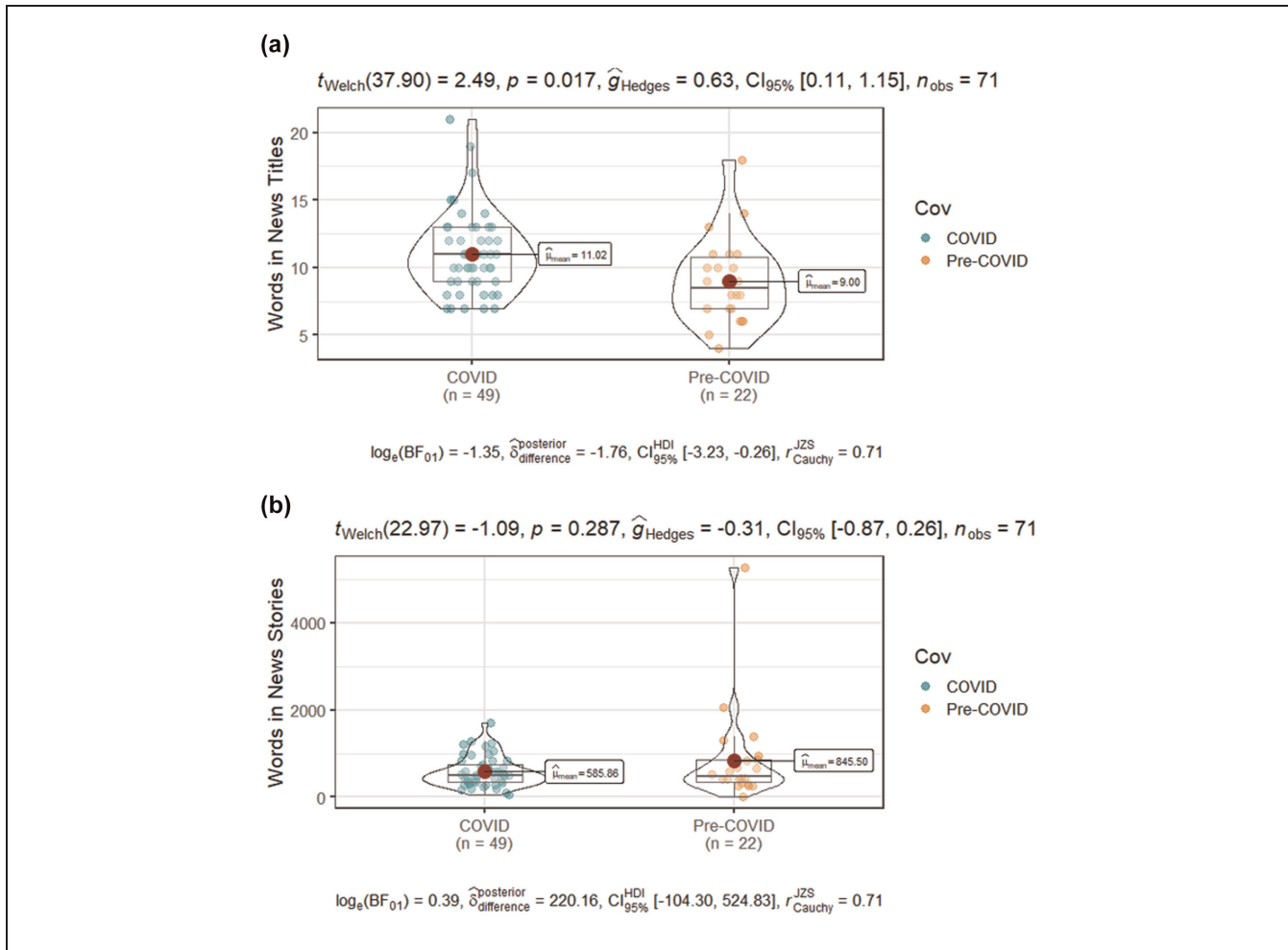
(a)

$t_{Welch}(37.90) = 2.49$, $p = 0.017$, $\widehat{g}_{Hedges} = 0.63$, $CI_{95\%}$ [0.11, 1.15], $n_{obs} = 71$

$\log_e(BF_{01}) = -1.35$, $\widehat{\delta}_{difference}^{posterior} = -1.76$, $CI_{95\%}^{HDI}$ [-3.23, -0.26], $r_{Cauchy}^{JZS} = 0.71$

(b)

$t_{Welch}(22.97) = -1.09$, $p = 0.287$, $\widehat{g}_{Hedges} = -0.31$, $CI_{95\%}$ [-0.87, 0.26], $n_{obs} = 71$

$\log_e(BF_{01}) = 0.39$, $\widehat{\delta}_{difference}^{posterior} = 220.16$, $CI_{95\%}^{HDI}$ [-104.30, 524.83], $r_{Cauchy}^{JZS} = 0.71$

Figure 1. Text word counts in news titles and news stories: (*a*) comparison between words in news titles and (*b*) comparison between words in news stories.

- Step 3. Network analysis: Network analysis involves the explanation of the nodes and edges based on the network plot. Usually, a large node indicates the presence of a keyword, which is associated with many other words with some specific high edge association with particular words in the form of higher co-occurrence in the corpus. The presence of such nodes can be considered as clusters or key topics determined by the word with large node and higher edges for different combination of word pairs.

In this study, the news articles are divided into two groups for text network analysis. The following section explains the results of text network analysis.

News Articles During COVID-19

Figures 2 and 3 show the text networks from news articles from before and during the COVID-19 pandemic.

Figure 2 shows the text network for news articles published during the pandemic. The top features include
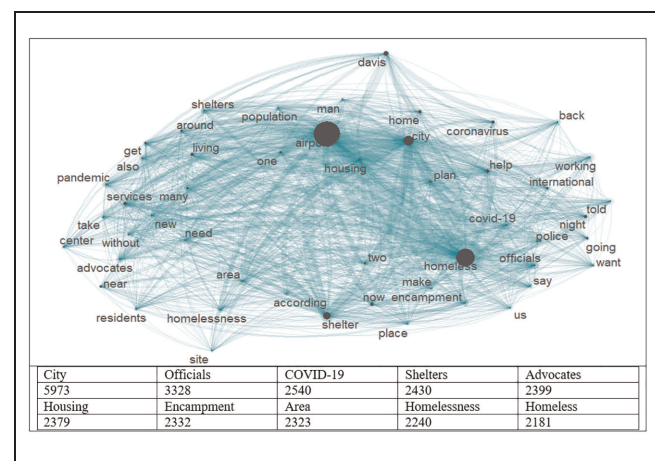

City	Officials	COVID-19	Shelters	Advocates
5973	3328	2540	2430	2399
Housing	Encampment	Area	Homelessness	Homeless
2379	2332	2323	2240	2181

Figure 2. Text network from news media during COVID-19 pandemic.

"city," "officials," "COVID-19," "shelters," "advocates," "housing," "encampment," "area," "homelessness," and "homeless." Visually, there do not appear to be many words in the exact center of the network. However, the words that appear to be closest include "housing," "two," "homeless," "COVID-19," "one," "area," and "according." The words "Davis," "back," "told," and "site" appear to be the furthest away from the center. The words with the largest nodes are "airport," "homeless," "shelter," "covid," "coronavirus," and "city" and there are clusters surrounding these nodes. "Shelter" and "encampment" are widely discussed in these news articles.

Collocation indicates a set of words occurring together in a text corpus (i.e., a document with texts). Collocation helps to understand unambiguous meaning or connotation for word groups from the unstructured dataset by extracting information. This study extracts collocation in the form of rigid collocations by extracting n-grams (in this study, bigram and trigram means group of two or three words together) that always occur side by side by keeping the same order. Table 2 shows the top 15 collocation measures for the COVID-19 text corpus, organized by count, length, lambda, and z-value. "Experiencing homelessness" is the first collocation and has the highest z-value of 12.549. Lambda identifies the probability that exactly the mentioned two words appear in the text next each other. Z is also a z-standardized lambda. "Baggage claim" has the highest lambda value (9.439), and "international airport" has the highest count. All the collocations have the same length, and the z-values are all fairly like each other (with the highest z-value being 12.549 and the lowest being 10.227). It is worth mentioning that the "mental health" issue has been recognized by the news media (the term shows an occurrence of eight times with z-value of 11.246). The collocation analysis identified some of the key characteristics such as "new homelessness experience," "mental health," "social security," "unsheltered population," and "outreach workers."

News Articles Before COVID-19

Figure 3 shows the text network for pre-COVID. "Children," "Atlanta," "homeless," "percent," "families," "number," "many," "homelessness," "house," and "national" appear as the top features. The words "new," "shelters," "home," "years," and "way" visually appear to be closest to the center of the text network, whereas "terminal," "travelers," and "officials" appear to be the furthest from the center. Compared with Figure 2, Figure 3 has words much closer to the center of the network. It also seems to have a greater focus on family, including words such as "children" and "families" in its key features. Figure 2, on the other hand, seems to have
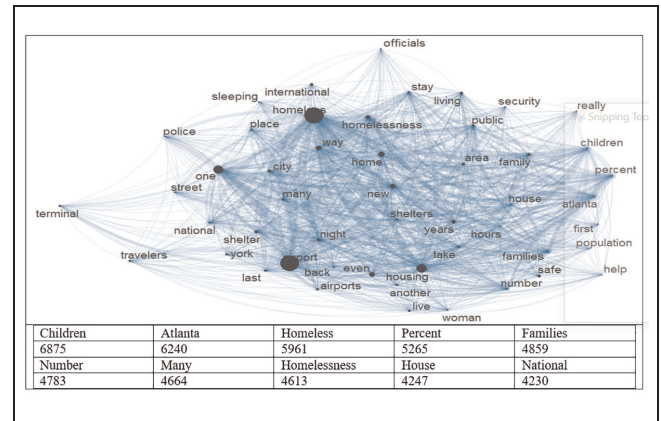

Children	Atlanta	Homeless	Percent	Families
6875	6240	5961	5265	4859
Number	Many	Homelessness	House	National
4783	4664	4613	4247	4230

Figure 3. Text network from news media before COVID-19.

more of a focus on words such as "officials" and "advocates." Figure 3 also seems to have more of a focus on numbers, including "percent" and "number" as key features. The two figures share some features, however, including "homeless" and "homelessness," and although the other key features do not directly repeat, they have similar themes (such as location and a focus on shelters). The words with the largest nodes are "one," "homeless," "airport," and "housing." The pre-COVID text network shows clusters surrounding these nodes.

The top 15 collocation measures for the pre-COVID text corpus are shown in Table 3, organized by count, length, lambda, and z-value. "Reagan national" has the highest z-value (12.650), "minimum wage" has the highest lambda, and "New York" has the highest count (n = 23). All the collocations have the same length. It is interesting to note that the z-values had a greater change in this table when compared with Figure 2. In Figure 2, the highest z-value is 12.549 and the lowest is 10.227 (a difference of 2.322), whereas in Figure 3 the highest z-value is 12.650 and the lowest is 6.567 (a difference of 6.083). "Minimum wage" has been listed in the top 15 collocation word list. The city names (New York, Washington D.C., and San Francisco) in the list indicate the airports that were more affected in pre-COVID period. It is well known that housing prices in these cities are extremely high. Compared with COVID-19 collocation words, pre-COVID collocation words do not provide enough evidence on the characteristics of the homeless population. Mental health issues were not present in the top collocation words in pre-COVID data. Some of the potential solutions in the top collocation words are "new policy," and "safe place." It is unclear whether and to what extent safe living space, food, and health insurance can be provided among homeless persons. There is no universal solution to address homelessness and news media most often describe the issue rather than providing any solution.

Table 2. Collocation Measures of Bigrams Using COVID Text Corpus

No.	Collocation	Count	Length	Lambda	z
1	Experiencing homelessness	15	2	7.817	12.549
2	International airport	23	2	5.221	12.306
3	Philadelphia international	7	2	5.861	11.722
4	Two years	7	2	5.451	11.480
5	Mental health	8	2	7.632	11.246
6	City officials	15	2	3.434	11.207
7	Social security	6	2	6.107	11.188
8	Baggage claim	9	2	9.439	11.066
9	Providers advocates	6	2	6.306	10.982
10	Unsheltered population	6	2	5.714	10.933
11	Coronavirus pandemic	7	2	4.831	10.775
12	Outreach workers	5	2	6.066	10.568
13	Action news	5	2	7.283	10.467
14	COVID-19 pandemic	7	2	4.538	10.372
15	Federal aviation	5	2	6.946	10.227

Table 3. Collocation Measures of Bigrams Using Pre-COVID Text Corpus

No.	Collocation	Count	Length	Lambda	z
1	Reagan national	13	2	8.017	12.650
2	International airport	17	2	4.935	11.121
3	Boarding pass	6	2	7.834	10.487
4	National airport	14	2	4.603	10.480
5	New policy	6	2	5.162	9.297
6	Safe place	5	2	5.070	9.261
7	Every night	5	2	4.805	8.909
8	Washington DC	5	2	8.599	8.588
9	Homeless population	7	2	3.687	7.740
10	Airport officials	7	2	3.521	7.517
11	Number homeless	7	2	3.319	7.412
12	Sleeping airport	6	2	3.366	6.873
13	New York	23	2	9.599	6.663
14	San Francisco	6	2	10.232	6.638
15	Minimum wage	5	2	10.317	6.567

Airport Interventions to Handle Homelessness

To tackle the homelessness issue, airport authorities work with local city governments, the Federal Aviation Authority (FAA), and other involved agencies involved in executing programs that help relocate homeless people from the airport. This could also include supporting government and non-government authorities to help relocate the homeless people to transitional shelters or to connect them with their relatives (*24*). There are examples in which airport authorities have taken the homelessness issue even more seriously and they have been found to be actively working with homeless people. Airports often employ different long-term intrusion approaches for facilitating treatment, minimizing violence, and reducing recurrent encounters with the first responders by coupling law enforcement with other needed workforce such as psychologists, nurses, and social workers For example,

a crisis intervention team has been employed in Atlanta Airport in Georgia that helped deescalate violent or aggressive situations and provided mental support to the people involved, including largely homeless people (*25*). During COVID-19, some airports may have been taking stricter measure to ensure the safety of people. For example, the Philadelphia airport authority worked with the city and helped transfer homeless people in the airport to the quarantine rooms with shelter and treatment beds.

Conclusion

The topic "airport and homelessness" has become one of the critical issues in the aviation sector. News media have extensively featured this issue. However, quantification and content analysis of these featured news have not been explored before. In recent years, several studies

have explored news media and social media data to identify transportation-related insights (*26–32*). To mitigate this gap, this study performed a thorough literature review and news media mining by collecting relevant news articles on this issue to answer three major research questions.

The study answered these questions by performing text network analysis and collocation analysis with the inclusion of relevant news media articles. The COVID-19 text network reveals topics such as "shelter" and "mental health." The pre-COVID text network shows topics such as "minimum wage" and "housing." Some cities are highlighted in the collocation list in the pre-COVID period. For the COVID-19 pandemic list, only Philadelphia is listed, which indicates that "airport and homelessness" is associated with most of the airports throughout the U.S. The pre-COVID text network indicates that the listed cities (New York, Washington D.C., and San Francisco) in the collocation are associated with high housing prices. Pre-COVID news articles provide contexts of children, wages, and families, and focus more on "city," "shelters," "encampment," and "advocates." The text network analysis shows that news coverage changed significantly between these two time periods, both in quantity and topic coverage. Although some of the keywords indicate that solutions are also discussed, these are limited to shelter and housing related issues. It is also worth mentioning that mental health, wages, and children related issues were also introduced. Social support to the homeless population can be considered as the measures and interactions by which one affluent individual or group directly provides the homeless population with resources, needs, and a sense of affirmation. Structural social support is often mentioned in the news media stories; however, the media coverage is often limited to mentioning the issue or problem rather than providing any solution. The current study proves that news media mining can be used as a supportive approach to the conventional methods in exploring this critical issue. The contribution of this study is that it represents an advancement of a first application of text mining to extract information from unstructured news articles.

The number of homeless individuals has critically increased in the U.S. during the last few years. Many factors have exacerbated the homelessness issue, such as housing problems and increased rent levels after the beginning of the COVID-19 pandemic, and inadequate shelter. The news reporting shows that increased numbers of homeless individuals were attracted to airports, more than before the pandemic, because of advantages such as safe environment and convenient access to food and necessary amenities. However, airports are required to allocate extra resources (i.e., staff and budget) to manage the presence of homeless people within their boundaries while respecting the individuals who are experiencing this situation. On the other hand, travelers usually do not expect to see homeless individuals within the airport and may demand a response to the issue. Homeless individuals, because of the lack of health care opportunities and considerably low quality of life, may have physical or mental challenges and consequently might add specific safety and security concerns at airports. More research is needed to assist airports in developing and implementing policies and strategies to address homelessness at their facilities.

The results of this study need to be interpreted in the light of the following limitations. First, the data collection is limited. Seventy-one relevant articles were collected by web scraping from Google News Alert-generated URLs. Future studies should consider the inclusion of more news articles on the topic. Additionally, mining social media on similar topics can provide additional insights. Second, the current analysis uses a descriptive text mining tool known as text network which has advantages in unearthing important insights. Nonetheless, there are several areas where future research could provide more advanced insights into the use of other advanced NLP techniques, such as topic modeling, on this topic. The first is to capitalize on the capability of narratives to represent the current state of safety, whereas the fixed fields are locked into the knowledge available at the time of database design. This study's findings can be validated by conducting interviews with airport authorities and homeless individuals to gain a deeper understanding of the issues and potential solutions.

Author Contributions

Declaration of Conflicting Interests

Funding

ORCID iDs

Subasish Das https://orcid.org/0000-0002-1671-2753
Javad J. C. Aman https://orcid.org/0000-0001-5656-4698

M. Ashifur Rahman (iD) https://orcid.org/0000-0001-6940-1599

References

1. Schwartz, R. The End of the Line: The Homeless and the Transportation Industry. *Portfolio*, vol.2, 1989, pp. 38–46.
2. Berendt, B. Text Mining for News and Blogs Analysis. In *Encyclopedia of Machine Learning* (C. Sammut, and G. I. Webb, eds.), Springer, Berlin, Heidelberg, 2010, pp. 968–972.
3. Bassett, E., A. Tremoulet, and A. Moe. *Relocation of Homeless People From ODOT Rights-of-Way*. No. OTREC-RR-12-14. Transportation Research and Education Center (TREC), Portland, OR, 2013.
4. Hui, V., and K. N. Habib. Homelessness vis-à-vis Transportation-Induced Social Exclusion: An Econometric Investigation of Travel Behavior of Homeless Individuals in Toronto, Canada. *Transportation Research Record: Journal of the Transportation Research Board*, 2017. 2665: 60–68.
5. Boyle, D. K. *Transit Agency Practices in Interacting With People Who Are Homeless: New Possibilities*. Transportation Research Board, Washington, D.C., 2016, p. 17.
6. Hui, V., and K. N. Habib. *An Investigation of Transport-Related Social Exclusion of the At-Risk Community (Homeless People) in Toronto, Canada*. Transportation Research Board, Washington, D.C., 2014, p. 18.
7. Lamers, J., S. Farber, and K. N. Habib. *Accessibility to Community Services in the City of Toronto: A Perspective on Social Disadvantage and Homelessness*. Transportation Research Board, Washington, D.C., 2016, p. 20.
8. Grimes, A., and M. Smirnova. Perspectives on an Earn-a-Bike Intervention on Transportation, Health and Self-Esteem Among Men Experiencing Homelessness. *Journal of Transport & Health*, vol.18, 2020, p. 100904.
9. Cronley, C., M. Huslage, and E. M. Roark. *Testing the Psychometric Properties of a Scale to Measure Transportation Access Among Individuals Experiencing Homelessness*. Transportation Research Board, Washington, D.C., 2021, p. 15.
10. Jocoy, C. L., and V. J Del Casino.. *The Mobility of Homeless People and Their Use of Public Transit in Long Beach, California*. University of Southern California, Los Angeles, 2008.
11. Jocoy, C. L., and V. J. Del Casino. Homelessness, Travel Behaviour, and the Politics of Transportation Mobilities in Long Beach, California. *Environment and Planning A*, vol.42, 2010, pp. 1943–1963.
12. Loukaitou-Sideris, A., J. L. Wasserman, R. Caro, and H. Ding. *Homelessness in Transit Environments Volume I: Findings From a Survey of Public Transit Operators*. Institute of Transportation Studies, University of California, Los Angeles, 2020.
13. Loukaitou-Sideris, A., J. L. Wasserman, R. Caro, and H. Ding. *Homelessness in Transit Environments Volume II: Transit Agency Strategies and Responses*. Institute of Transportation Studies, University of California, Los Angeles, 2021.
14. Rudy, J., and A. J. Delgado. Homelessness and Problems it Presents to the Orange County Transportation Authority "A Social Approach to a Social Problem." *Proc., 2006 Bus and Paratransit Conference*, Orange County, CA, 2006, p. 3.
15. Carlson, D. L., S. Reder, N. Jones, and A. Lee. *Homeless Student Transportation Project Evaluation*. Washington State Transportation Center (TRAC), Seattle, 2006.
16. Bassett, E. M., and A Tremoulet.. *Addressing Homeless Encampments on DOT Land: The Case of Oregon's Baldock Rest Area*. Transportation Research Board, Washington, D.C., 2013, p. 18.
17. Berendt, B. Text Mining for News and Blogs Analysis. In *Encyclopedia of Machine Learning* (C. Sammut, and G. I. Webb, eds.), Springer, Berlin, Heidelberg, 2016, pp. 1–9.
18. de Fortuny, E. J., T. De Smedt, D. Martens, and W. Daelemans. Media Coverage in Times of Political Crisis: A Text Mining Approach. *Expert Systems With Applications*, vol.39, No. 14, 2012, pp. 11616–11622.
19. Fayyad, U., and G Piatetsky-Shapiro.. From Data Mining to Knowledge Discovery in Databases. *AI Magazine*, vol.17, 1996, pp. 37–54
20. Aman, J. J., J. Smith-Colin, and W. Zhang. Listen to E-Scooter Riders: Mining Rider Satisfaction Factors From App Store Reviews. *Transportation Research Part D: Transport and Environment*, vol.95, 2021, p. 102856.
21. Suh, J. H., C. H. Park, and S. H. Jeon. Applying Text and Data Mining Techniques to Forecasting the Trend of Petitions Filed to E-People. *Expert Systems With Applications*, vol.37, 2010, pp. 7255–7268.
22. Schumaker, R. P., and H. Chen. Textual Analysis of Stock Market Prediction Using Breaking Financial News. *ACM Transactions on Information Systems*, vol.27, 2009, pp. 1–19.
23. Aman, J. J., and J. Smith-Colin. Leveraging Social Media to Understand Public Perceptions Toward Micromobility Policies: The Dallas Scooter Ban Case. *Findings*, 2021, p. 21328. https://doi.org/10.32866/001c.21328.
24. Herring, C. The New Logics of Homeless Seclusion: Homeless Encampments in America's West Coast Cities. *City & Community*, vol.13, No. 4, 2014, pp. 285–309.
25. McGriff, J. A., B. Broussard, B. N. D. Neubert, N. J. Thompson, and M. T. Compton. Implementing a Crisis Intervention Team (CIT) Police Presence in a Large International Airport Setting. *Journal of Police Crisis Negotiations*, vol.10, No. 1–2, 2010, pp. 153–165.
26. Das, S. Fatal Crash Reporting in Media: A Case Study on Bangladesh. *Transportation Research Record: Journal of the Transportation Research Board*, 2021. 2675: 960–971.
27. Brooke, K., M. Thomson, C. Johnson, S. Cacsl, C. Pirkle, and T. Sentell. Public Health Framing in Local Media Coverage of Crashes Involving Pedestrians or Bicyclists in Hawaii 2019: A Content Analysis. *Transportation Research Interdisciplinary Perspective*, vol.13, 2022, p. 100525.
28. Das, S., and H. Zubaidi. City Transit Rider Tweets: Understanding Sentiments and Politeness. *Journal of Urban Technology*, 2021, pp. 1–16. https://doi.org/10.1080/10630732.2021.1903288.

29. Dutta, A., and S. Das. Tweets About Self Driving Cars: Deep Sentiment Analysis Using Long Short-Term Memory Network (LSTM). In *Proc., Advances in Intelligent Systems and Computing: International Conference on Innovative Computing and Communications* (D. Gupta, A. Khanna, S. Bhattacharyya, A. E. Hassanien, S. Anand, and A. Jaiswal, eds.), New Delhi, India, February 20–21, 2021, Springer, Singapore, pp. 515–523.

30. Hong, Y., Q. Ma, Z. Wang, Q. Cai, K. Xie, and D. Yang. Safety of Micro-Mobility: Analysis of E-Scooter Crashes by Mining News Reports. *Accident Analysis & Prevention*, vol.143, 2020, p. 105608.

31. Das, S., A. Dutta, G. Medina, L. Minjares-Kyle, and Z. Elgart. Extracting Patterns From Twitter to Promote Biking. *IATSS Research*, vol.43, No. 1, 2018, pp. 51–59.

32. Das, S., A. Dutta, T. Lindheimer, M. Jalayer, and Z Elgart. YouTube as a Source of Information in Understanding Autonomous Vehicle Consumers: A Natural Language Processing (NLP) Study. *Transportation Research Record: Journal of the Transportation Research Board*, 2017. 2673: 242–253.