

# Exploration of social media data to understand the perceived accessibility in transportation infrastructure

Steven Tanner McCullough  
steven.mccullough@mavs.uta.edu  
Department of Civil Engineering  
The University of Texas at Arlington

June Young Park  
june.park@uta.edu  
Department of Civil Engineering  
The University of Texas at Arlington

## ABSTRACT

This paper explores the potential of social media data to evaluate airport accessibility for passengers with disabilities (PWD). We utilized geo-spatial information and user-generated content to identify the disability related comments from 64 hub airports in the US, focusing on the perceived accessibility issues from PWD. In total, we found 321,858 reviews with 6,862 of these reviews relating to accessibility related issues. Reviews from PWD show lower satisfaction score than the overall airport passenger reviews. Our findings indicate that accessibility challenges could be specifically related to the layout and operation of airport facilities rather than airport size and/or overcrowd. There will be multiple future research directions for understanding and enhancing accessibility in urban infrastructure through this innovative use of publicly available data.

## CCS CONCEPTS

• Applied computing; • Human-centered computing;

## KEYWORDS

social media, data analytic, accessibility, zero-shot classification, transportation infrastructure

## 1 INTRODUCTION

According to the 2020 Bureau of Transportation Statistics, an estimated 926 million travelers flew by air in the United States, 27 million of these travelers have a reported disability. The number of disability related complaints to the US Department of Transportation increased by 54% in 2021 from pre-covid levels in 2019. [6] Analyzing such traveler feedback has derived meaningful outcomes to comprehensively understand the performance of airports for customer satisfaction [4]. Conventionally, those feedback data collections were conducted by airport customer satisfaction team with digital surveying or in situ interviews with passengers with disability (PWD). However, such methods can be plagued by survey fatigue, poor response rate, and other shortcomings of mass surveying. Additionally, airports are not the ones who have contact information for many travellers; airlines are the ones who collect personal information during the ticketing process. This split

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

*UrbanAccess '24, October 24, 2024, Virtual*

© 2024 Association for Computing Machinery.

functionality of airline and airport results in a deficiency in data collection process [2].

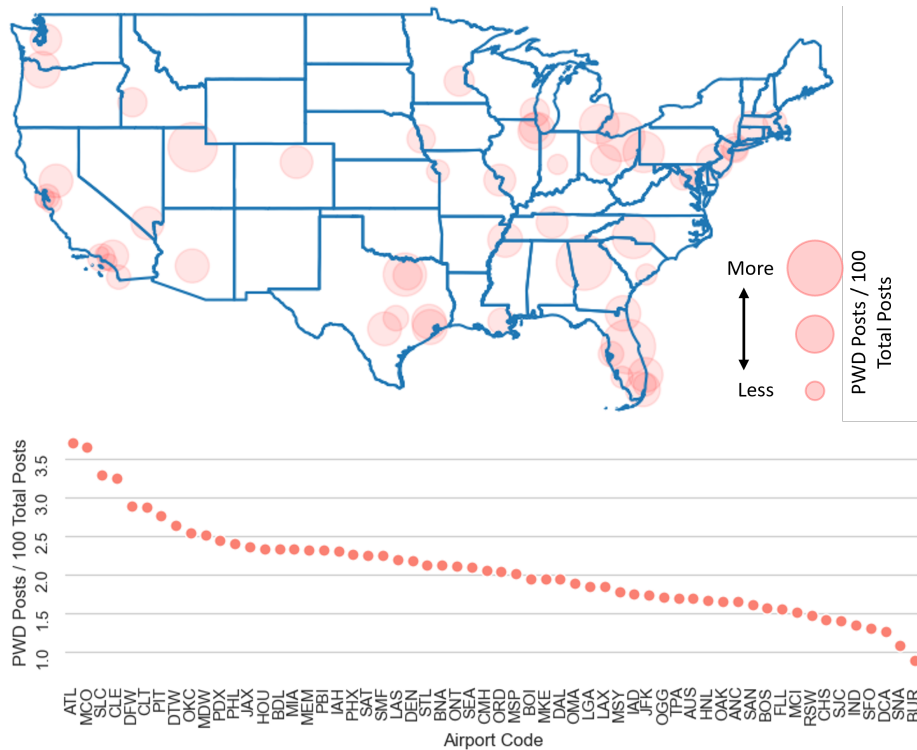
Ubiquitous computing and social media have made it easy for citizens to post their perceptions in the built environment. With advanced text mining techniques, such data collections have provided critical insights for facility managers to understand occupant satisfaction in urban infrastructure (e.g., transportation) [8]. The main motivation for leaving reviews on social media is to warn other travelers and notify facilities management for them to rectify the issue [9]. Gitto and Macuso found strengths in using social media to improve airport services through analyzing sentiment analysis of posts on social media [3]. Park et al. were able to mine Google Maps reviews to find the stressors that passengers encountered in airports during the Covid-19 pandemic [7].

Similarly, this paper explores the potential of social media data to investigate the perceived accessibility of PWD at airports. Section 2 explains data collection process as well as filtering criteria for specific disability related posts using zero-shot classification. Then, we conduct exploratory data analysis and highlight notable individual reviews in Section 3. Finally, Section 4 concludes the paper with potential research directions by using the explored data.

## 2 DATA COLLECTION AND PRE-PROCESSING

We collected textual reviews for airport businesses on a major mapping and navigation platform, Google Maps, for 64 airports. Individual airports were selected based on Federal Aviation Administration's designations for airport size (i.e., P-L: 30 large hub airports, P-M: 34 medium hub airports). These designations were chosen as they include the largest airports in the US, which collectively account for approximately 87% of total passenger air traffic in the US as of 2021. The data source (Google Maps) is one of the most widely used map services, with over 1 billion daily active users reported. It holds a significant market share in navigation services and offers extensive data, including hundreds of thousands of reviews for major US airports. Each review includes a rating from 1 to 5, along with a text description, photos and recommendations.

Reviews were collected using data scraping techniques, but did so in such a way that did not request continuous data to run the webpage, the scripts used only requested the review text from the Google Maps database. This method proved more reliable and efficient than traditional scraping but still provided less precise metadata (e.g., dates are approximate). Each request retrieves 10 reviews and is sent every roughly every 5 seconds, capturing text, ratings, and estimated dates (e.g., 1 year ago, 6 months ago). This method was implemented to navigate the limitations of data scraping information that is controlled by JavaScript.



**Figure 1: Top: A spatial mapping of US airports and the ratio of disability related complaints. Larger circles indicate more complaints per total passengers. Bottom: A strip plot of all of the mapped airports as shown above but with the actual ratios of PWD comments per 100 total comments.**

It is important to note that the data collection itself contain all the customer reviews on 64 airports. To classify what reviews were posted by an individual that had a disability, we used zero-shot classification due to it’s ability to work with data specific training [1]. This classification method used the BART model to allow for easy input of candidate labels and effective performance in classifying reviews based on their text. Reviews were classified according to three different labels: 1) User had a disability, 2) User did not have a disability, and 3) User disability status unknown. Three labels were added to allow for the model to have a higher level of confidence in tagging reviews. Only reviews with a self quantified confidence score of 75% or higher were considered. Lastly, we only select the posts at least more than 20 characters to extract the meaning opinion of the perceived accessibility.

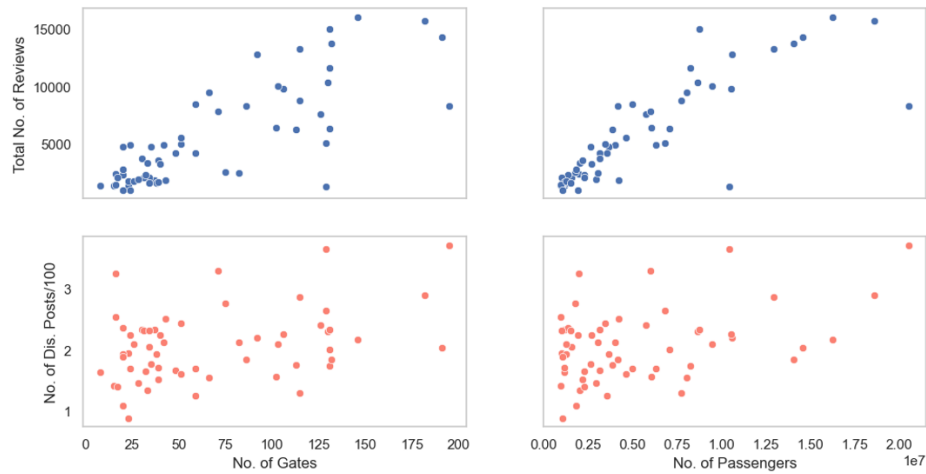
**Table 1: Data collection summary**

Airports	Total Posts	Posts per airport	PWD Posts	GMap Score	GMap Score (PWD)
30 P-L	248,271	8,275	5,431	3.75	3.10
34 P-M	73,587	2,164	1,431	4.09	3.58

### 3 EXPLORATORY DATA ANALYSIS

Table 1 summarizes our data collection for further analysis. The ratio of posts collected lie mostly with the large hub airports with more than 77% of total posts collected, the ratio of disability related complaints correlates quite similarly with slightly more than 78% of the complaints also stemming from large hubs. P-L airports indicate lower Google Maps score than their P-M counterparts overall and even more so with travellers with disabilities with a rating of only 3.1 stars. However, PWD score even lower than the overall score left by all travellers by nearly half a star for both large and medium hub airports, indicating that many of the PWD leaving comments not only score lower but likely have encountered barriers to their travel more than all passengers would. Figure 1 displays the visual representation of the ratio of disability related reviews for each airport across the nation. The lower half of figure 1 shows how many disability related posts occur for every 100 total posts. These ratios range from 1 to 3.5 posts for every 100 posts with the majority of airports residing under a ratio of 2.5%. Geographically, there is not a direct correlation to larger or lesser populated areas when examining the rate of disability related reviews.

Given that PWD comments are generally more negative, there can be some assumption that airport size and overcrowd are likely to blame. However, in figure 2, there is only correlation shown between an airports size and passenger volume and its total number of overall reviews (blue). When normalizing the posts to show



**Figure 2: Airport size (by number of gates) and crowd (by passenger volume) plotted against total reviews (blue) and prevalence of disability reviews (red)**

the rate of PWD complaints per 100 total posts, there is no link between the number of gates or the number of passengers flowing through and the ratio of disability related posts left for each airport (red). Again, given that these posts are more negative than their overall counterparts, it could be assumed that there is an issue that impedes accessibility that is not related to overcrowd but rather an implication of building design or operations. Similarly, there does not appear to be a correlation with the occurrence of disability related comments (praise or concern) when examining the level of passenger volume or the size of the airport, measured as the total number of gates. While the rate of comments and disability related comments do increase with passenger volume, the ratio of disability comments varies and does not correlate with these descriptors indicating that something else is the causation of these comments.

### 3.1 Notable individual reviews

We highlight notable individual reviews to provide an anecdotal insight on how PWD perceives the airport operations. Reviews from several airports are shown below, i.e.,

- (1) "It's a mess! If you don't fly Delta, you have to walk what seems like miles. Sign says b gates are a 7 minute walk...it's much longer with someone with special needs...With a handicapped Mother, that's torture. THE WORST AIRPORT I have ever visited, in the world, so far!!!", 1 Star GMap rating, Salt Lake City International Airport (SLC)
- (2) "Waiting on wheelchair assistance I almost missed my plane", 1 Star GMap rating, Hartsfield-Jackson Atlanta International Airport (ATL)
- (3) "Wheelchair access was good. Airport too big for me to navigate on my feet.", 4 Star GMap rating, Chicago O'Hare International Airport (ORD)
- (4) "No wheelchair assistance available. Bathrooms were disgusting. Impossible to maneuver wheelchair through TSA line.", 3 Star GMAP rating, William P. Hobby Airport (HOU)

These reviews illustrate the challenges faced by travelers with disabilities in various airports. The first quote expresses frustration with the long distances and inadequate signage which happens to occur during SLC's Terminal A expansion and revitalization plan. While one traveler from ATL describes the issues of relying on staff to navigate airports, this type of issues occurs in many airports. Not all travelers have caused for concern, a traveler at ORD found wheelchair assistance but also noted the difficulty of navigating large airport spaces, suggesting that size and sprawl of airports can be a limiting factor for many. Finally, while the last quote mentions cleanliness the traveler does mention the difficulty that they had when navigating the security checkpoint lines in HOU airport.

## 4 DISCUSSION

The methods and data used in this paper offer a powerful combination of real-time insights and community engagement to realize the perceived accessibility in transportation infrastructure. By using more advanced text mining techniques, researchers can analyze large amounts of data from posts, comments, and reviews to identify patterns, sentiment, and specific issues encountered by PWD. This approach does not only highlight problems in transportation infrastructure but also highlights the experiences of people that traditional surveys may overlook. The use of social media and review sites can circumvent the challenges of survey methods including survey fatigue and implementation hurdles as the data is freely exposed online rather than needing extraction from people. The first step in most survey techniques is initiating contact with the desired target population, be it email address collection, physical surveys or cold calling. However data mining social media and review sites remove this first step as these people have already initiated this first step by leaving their own experiences. Further research is being performed to analyze the specific areas in airports that give the most issues to PWD as well as how they relate to current legislature and design standards. Further future work can include total mapping of accessible areas though analyzing social media with advanced text mining, returning more specific insights

like the specific cause of inaccessibility and if it is currently covered under legislation or not. Social media platforms are a repository of user-generated content, where individuals share their experiences, challenges, and suggestions related to accessibility. This is indeed valid for citizens with disabilities to facilitate a more participatory approach to urban planning and development [5]. By fusing multiple data sources, cities can create a more inclusive environment, ultimately leading to enhanced mobility and quality of life for all inhabitants.

However, the method presents a few challenges with how to tackle bias as not every traveler is required to complete a review on social media. For instance, a sizeable portion of reviews are performed on behalf of another person that does have a disability rather than the poster having it themselves, those with the disabilities are either unable or do not wish to post the review themselves. Under coverage of these issues is a potential issue as well as the poster has to specifically refer the disability being a part of the experience, regardless of sentiment.

## ACKNOWLEDGMENTS

This work was supported by a grant from the Center for Transportation Equity, Decisions, and Dollars funded by the US Department of Transportation Research and Innovative Technology Administration (OST-R).

## REFERENCES

- [1] Waad Alhoshan, Alessio Ferrari, and Liping Zhao. 2023. Zero-shot learning for requirements classification: An exploratory study. *Information and Software Technology* 159 (7 2023), 107202. <https://doi.org/10.1016/j.infsof.2023.107202>
- [2] Jerusa Barbosa Guarda de Souza, Talita Naiara Rossi da Silva, and Nilton Luiz Menegon. 2018. Air Travel Accessibility: Interaction Between Different Social Actors. In *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018)* (Florence). 45–54.
- [3] Simone Gitto and Paolo Mancuso. 2017. Improving airport services using sentiment analysis of the websites. *Tourism Management Perspectives* 22 (4 2017), 132–136. <https://doi.org/10.1016/j.tmp.2017.03.008>
- [4] D. Hutchinson, Vanja Bogicevic, Wan Yang, Anil Bilgihan, and Milos Bujisic. 2013. Airport service quality drivers of passenger satisfaction. *Tourism Review* 68 (10 2013), 3–18. Issue 4. <https://doi.org/10.1108/TR-09-2013-0047>
- [5] Steven McCullough, Jessica Eisma, June Young Park, Mikila Salazar, and Sarah F Rose. 2022. Inequity by inequity: community driven investigation of wheelchair user discomfort by infrastructure failures. In *Proceedings of the 9th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*. 274–277.
- [6] Government Accountability Office. 2021. PASSENGERS WITH DISABILITIES Airport Accessibility Barriers and Practices and DOT's Oversight of Airlines' Disability-Related Training Report to Congressional Committees United States Government Accountability Office. <https://www.gao.gov/assets/gao-21-354.pdf>
- [7] June Young Park, Evan Mistur, Donghwan Kim, Yunjeong Mo, and Richard Hoefler. 2022. Toward human-centric urban infrastructure: Text mining for social media data to identify the public perception of COVID-19 policy in transportation hubs. *Sustainable Cities and Society* 76 (1 2022). <https://doi.org/10.1016/j.scs.2021.103524>
- [8] Alireza Shamshiri, Kyeong Rok Ryu, and June Young Park. 2024. Text mining and natural language processing in construction. <https://doi.org/10.1016/j.autcon.2023.105200>
- [9] Sihyun Yoo, Bert van Wee, and Eric Molin. 2024. Long distance accessibility by air transportation: a literature review. *Transport Reviews* (2024). <https://doi.org/10.1080/01441647.2024.2322430>