Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company’s public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Toward human-centric urban infrastructure: Text mining for social media data to identify the public perception of COVID-19 policy in transportation hubs

June Young Park*, Evan Mistur, Donghwan Kim, Yunjeong Mo, Richard Hoefer

* Corresponding author.

ARTICLE INFO

Keywords:
COVID-19 policy
Public perception
Social media data
Text mining
Human-building interactions
Human-centric urban infrastructures

ABSTRACT

The COVID-19 pandemic has made transportation hubs vulnerable to public health risks. In response, policies using nonpharmaceutical interventions have been implemented, changing the way individuals interact within these facilities. However, the impact of building design and operation on policy efficacy is not fully discovered, making it critical to investigate how these policies are perceived and complied in different building spaces. Therefore, we investigate the spatial drivers of user perceptions and policy compliance in airports. Using text mining, we analyze 103,428 Google Maps reviews of 64 major hub airports in the US to identify representative topics of passenger concerns in airports (i.e., Staff, Shop, Space, and Service). Our results show that passengers express having positive experiences with Staff and Shop, but neutral or negative experiences with Service and Space, which indicates how building design has impacted policy compliance and the vulnerability of health crises. Furthermore, we discuss the actual review comments with respect to 1) spatial design and planning, 2) gate assignment and operation, 3) airport service policy, and 4) building maintenance, which will construct the foundational knowledge to improve the resilience of transportation hubs to future health crises.

1. Introduction

The Novel Coronavirus Disease (COVID-19) has significantly changed urban infrastructures and socioeconomic conditions in the United States. While the pandemic has impacted all sectors (Gude and Muire, 2021; Megahed and Ghoneim, 2020), transportation hubs, where people gather in large numbers when travelling between regions, remain especially vulnerable to public health concerns potentially resulting in the importation and exportation of the virus (Keil and Ali, 2011; Nakamura and Managi, 2020). In particular, airports create international health risks as they are often overcrowded due to a large volume of passengers (Edelson and Phypers, 2011; Gaskin et al., 2021; Ikonen et al., 2018). Thus, airports are recognized as key vectors in the spread of infectious diseases and have accelerated international transmission of several virus outbreaks including the COVID-19 pandemic (Browne et al., 2016; Kishore et al., 2020; Roy and Ghosh, 2020). While researchers have developed treatments and vaccines, implementation of the nonpharmaceutical interventions (NPIs) (e.g., wearing a mask, social distancing, and surface cleaning) recommended by the Centers for Disease Control and Prevention (CDC) was adopted as the primary COVID-19 policy to control the transmission of the virus in these spaces.

1.1. Public perceptions of COVID-19 policy

Public policies requiring implementation and individuals’ compliance with these NPIs are critical to ensuring they are appropriately applied and adhered to. Along with state and local health safety requirements, the Transportation Security Administration (TSA) has maintained an evolving set of NPI policies in order to mitigate risk of contagion, including requiring visitors to wear masks when indoors at airports, adjusting security checkpoints by maintaining greater distances between visitors and reducing contact with employees, and...
promoting visitors to follow social distancing guidance. Additionally, many individual airports implement their own policies prompting NPI compliance through behavioral interventions such as handing out free masks to visitors, placing stickers denoting appropriate distances to keep between people in queues, and installing signage reminding visitors to maintain safe health standards.

However, the public perceptions of the COVID-19 policy were not actively evaluated, which hinders airport operators to understand how passengers respond to such initiatives. While attention and compliance to CDC health safety recommendations is widespread, it is not universal. Belief in misinformation about COVID-19 (Bierwiaczonek et al., 2020) and the effectiveness of NPIs (Hornik et al., 2021) undermine compliance with COVID-19 health policies. Additionally, political ideology significantly drives individuals’ perceptions of risks from COVID-19 and their willingness to adhere to health policies (de Bruin et al., 2020). Noncompliance to NPIs such as mask wearing and social distancing is common in many settings, including US airports where masking requirements were often ignored during the pandemic (Elachola et al., 2020). Individuals’ compliance to the COVID-19 policy is highly heterogeneous in the US (Gadarian et al., 2021). Understanding how individuals perceive the pandemic and the initiatives implemented in response to it is critical to identifying how they will react, and comply, with policy.

Previous research has leveraged social media data to understand how citizens perceive public health interventions responding to the pandemic. Globally, posts and searches in social media (e.g., Twitter, Instagram, Google) related to COVID-19 increased significantly (Dyer and Kolic, 2020; Rotveta et al., 2020), providing critical insights of the emotional reaction created by COVID-19 policy (Ridhwan and Har- greaves, 2021). Data mining research using Twitter to conduct sentiment analysis increases understanding of public awareness regarding the COVID-19 and shows its potential applicability as a communication tool for health departments (Boon-Itt and Skunkan, 2020). Among various policies, social distancing was the most commented on and supported by Twitter users (Saleh et al., 2021).

The main objective of these recent studies was to identify the general perception of COVID-19 policies. It is important to note that implementing and ensuring compliance to such policies are critical for maintaining public safety for both passengers and employees in airports, as well as for securing the resilience of transportation infrastructures to health crises (Ali et al., 2020; Banholzer et al., 2020; Brauner et al., 2020). As COVID-19 policy has changed how people interact in buildings, understanding public perceptions is essential for successful implementations in different transportation hubs. Almost every US major airport has adopted the CDC guidelines. However, they currently lack efficient and accurate evaluation of the public perceptions regarding COVID-19 policy in airports. Thus, in this article, we investigate the spatial drivers of public compliance, which are compounded with human interactions in indoor spaces. Our results can inform future health policies aimed at changing public behavior by altering the structure of individuals’ choices, prompting them to make socially beneficial decisions voluntarily (Thaler and Sunstein, 2018).

1.2. Human-centric urban infrastructure

Increasing our knowledge about how individuals interact with COVID-19 policy in different spaces will enable architects and engineers to design and operate urban infrastructures more safely by tailoring them to the space in which they are implemented. The rapid developments of information technology sparked the notion of smart buildings and cities, which mainly utilize the knowledge of human interactions in cities. When urban infrastructures are equipped with sensor network and advanced data analytic (e.g., machine learning), such intelligent infrastructures can incorporate human related data to calculate the optimal operational policy for improving human experiences while minimizing environmental impacts (O’Brien et al., 2020; Park et al., 2019b), which eventually increases the satisfaction levels of various users in cities (Frontczak et al., 2012). The applications of the human-centric approach vary from the building controls of lighting (Park et al., 2019a) and HVAC (Gunay et al., 2018) as well as larger infrastructure systems (Dargin and Mostafavi, 2020).

1.2.1. Facility management using user feedback

Therefore, we argue that the facility management (FM) of transportation infrastructures should incorporate public perception of COVID-19 policy to provide a safer environment for occupants. FM refers to the process of integrating and aligning non-core services for business operation and maintenance; it coordinates diverse functions for operating, maintaining, and managing physical assets (Lavy et al., 2010; Mangano and De Marco, 2014; Tucker and Pitt, 2009). Dealing with current and potential complaints of users could help ongoing commissioning improve on reactive management of complaints (Goins and Moezzi, 2013). Occupants’ comments and complaints reveal critical and underevaluated areas in FM, where it evaluates the performance and suggests new services in the buildings. Large scale facilities (e.g., airports) are integrated with social, technical, and organizational systems, but some critical features of complaints are overlooked in a simple quantitative analysis of complaints (Goins and Moezzi, 2013; Leaman and Bordass, 2007). Thus, more integrated in-depth approaches are needed to fully understand the operative status of facilities and improve their functions for occupants.

As there are unprecedented amounts of data available regarding occupant feedback in large scale facilities, more advanced and efficient techniques have been emerging to deal with such datasets, including machine learning, natural language processing (NLP), pattern recognition, information retrieval database, and advanced visualization (Becerik-Gerber et al., 2012; Witten and Frank, 2002; Yan et al., 2020). Text mining of occupant’s feedback data is a particularly promising approach to manage complaints in the domain of FM and has been used to develop successful models in several previous studies.

Moezzi and Goins analyze post-occupancy evaluation (POE) surveys on indoor environment quality (IEQ) satisfaction of occupants in 192 large buildings (Moezzi and Goins, 2011). They examine open-ended text responses focusing on occupants’ opinions on the overall indoor environmental quality, including thermal and acoustic comforts. Dutta et al. used NLP-based methodology to extract operational strategies from occupant survey databases, employing lexicon-based sentiment analysis, binomial classification, frequency-based algorithm, association rule mining, and topic modeling to combine the textual data and maintenance complaint categories (Dutta et al., 2021). Among those methods, they used Latent Dirichlet Allocation (LDA) and Non-Negative Matrix Factorization for topic modeling. In addition, the text data of complaint logs were also used for FM. Mo et al. developed a model that reads service request texts posted by occupants and automatically assigns workforce and priority using the NLP technique (Mo et al., 2020). They used building maintenance records for three years from buildings on a university campus and employed machine learning algorithms. In addition to the solicited occupant feedback (POE surveys and complaint logs), Lee and Yu analyzed location-based service data (Google Maps reviews) to compare with the existing airport survey (Airport Service Quality program by Airport Council International) to demonstrate the applicability of user generated online reviews as an indicator of airport service performance (Lee and Yu, 2018). In fact, this study was conducted before the COVID-19 (2008 - 2016) and they identified 25 topics (e.g., accessibility, check-in, passport control, security, navigation, cleanliness, arrival services), as the main complaints from actual passengers.

1 https://www.tsa.gov/coronavirus
1.3. Contributions

In this article, we critically evaluate public perceptions of COVID-19 policy at major hub airports in the US. Methodologically, we used a short text clustering algorithm to discover a set of 103,428 reviews from the location-based social media platform, Google Maps. Although several existing studies investigate public perception of COVID-19 policy using text mining approach, our study is uniquely aimed to understand the spatial drivers of the policy compliance in specific buildings (airports). We also fuse different data sources (Google Maps rate, passenger statistics, and architectural features) to comprehensively evaluate public perceptions of COVID-19 policy. Furthermore, we discuss the potential improvements of FM in airports to increase the resilience of transportation infrastructures.

1.4. Organization

The reminder of the article is organized as follows. In Section 2, we explain the methodology of our data analysis, which includes data collection, topic modeling, and data-driven discovery. Next, Section 3 summarizes the results of our study. Section 4 discusses recommendations for the airport management during the COVID-19 pandemic and future health crises. Finally, we conclude the article in Section 5.
2. Methodology

To identify the public perception of COVID-19 policy at airports, we employ a data driven approach, which consists of three major steps (Fig. 1). As the first step, we collect text data via Google Maps reviews from the population of large and medium hub airports in the US. Compared to other social media data, Google Maps platform not only provides review text data but also associated information (e.g., rate). More importantly, these location-based review data are specifically retrieved from certain airports, where we can further integrate other information (e.g., passenger statistics, regional and architectural features). To analyze our data collection efficiently, we used a topic modeling technique to cluster individual reviews into a few representative topics. Lastly, we identify each topic discussed in the reviews to understand the public perceptions at airports and further suggest the human-centric design and operation of airports.

2.1. Airport selection and regional information

The Federal Aviation Administration (FAA) categorizes airports into large hub (receives 1% or more of the annual US commercial emplanements), medium hub (receives 0.25 - 1% of the annual US commercial emplanements), small hub (receives 0.05 - 0.25% of the annual US commercial emplanements), and nonhub (receives less than 0.05% but more than 10,000 of the annual US commercial emplanements) types. Based on the numbers of passengers in 2020, FAA categorized 29 large hub airports and 35 medium hub airports. The total volume of passengers from both large and medium hub airports (64 airports in total) is 325,146,062 passengers, which accounts 88% of the total numbers of passengers in the US airports. Examining the population of medium and large hub airports in our analysis provides a representative look at major airport transportation hubs.

In addition to the hub types, we also classified airports by their terminal configurations. Fig. 2 illustrates different configurations by four basic types: linear, pier, satellite, remote hardstand (Federal Aviation Administration, 2018). Linear airports consist of a long-shaped main terminal that aircraft parked perpendicularly; easy access is one of its main advantages. Airports with a pier configuration have piers extending from the terminal processor where both sides can be used for passenger gates, providing efficient space utilization. However, Pier airports may have conflicts between arriving and departing passenger movements and cause longer passenger walking distances. Satellite terminal configurations are physically separated from the main terminal processor, and require a connection method (e.g., tunnels, bridges, underground walkways, moving walkways, or automated people movers). This configuration requires higher operating and maintenance costs.

2 https://www.faa.gov/airports/planning_capacity/categories/
Airports with a remote configuration is uncommon in the US. Figs. 3 & 4 summarize the meta-information of our airport selection. The airports in our sample are located across the nation in 35 states. The majority of the airports are located in South Atlantic and Pacific census region. California has the most airports (9 airports) in our dataset, but Florida and Texas follow with 7 and 6 airports respectively; other states have 1 or 2 airports. Almost half of the airports we study are classified as pier type airports and vary in terms of the numbers of passengers served. We collect both the numbers of gates and passengers as the number of gates is highly correlated with the number of passengers. As shown in Figure 4c, there are 6 airports (i.e., ATL, DFW, DEN, ORD, LAX, CLT), which serve the most passengers. Details on these data is tabulated (Table 4) in the appendix.

Lastly, we collect political data describing state legislative ideology and archival web data detailing regional COVID-19 policies that were in use during the data collection period. Political data is drawn from the state ideology dataset (Berry et al., 2010) and provides detailed information about the level of conservatism or liberalism in state politics along an index of liberalism from 0 to 100. Of the states in our dataset, South Carolina is the most conservative with an index of 17.8 and California is the most liberal with an index of 70.4. The additional perspective allows us to identify key policy changes over time, helping us account for the dynamic policy landscape airports are embedded within when interpreting our findings. This provides a contextual basis for understanding events over the period of our data and identifying basic correlations between the changing airport use, policy landscape, and passengers’ review content. We focus on four main state-level policies relevant to airport use: states of emergency, stay-at-home orders, mask mandates, and quarantines. These policies had the largest impact on air transit during this period, offering a representative impression of how critical state-level policies unfolded.

2.2. Review data collection

Considering that the first cases in the US were reported in January 2020, we collected all Google Maps review data over a period of 17 months from January, 2020 to May, 2021. Although the pandemic situation is still continuous, the month of May 2021 was remarkable with respect to the vaccination rate (at least 1 dose: 51.1% and fully vaccinated: 41.2%). We collected the text data by using the Outscraper, which is an API tool to scrape the social media data. Each review on specific airports contains both text contents and rate (1 - 5 stars) of user
experiences. Thus, we collected three information (i.e., review text, posted date, and rate) of Google Maps review data to analyze the public perception of COVID-19 policy.

We then filtered each review by the COVID-19 related topics, referring to the COVID-19 related keywords from previous literature (Chen et al., 2020; Kurian et al., 2020). Table 1 summarizes the keywords selection, which was case-intensive and combined by OR logical operation. Lastly, we selected only English reviews for the next topic modeling step.

### 2.3. Topic modeling

The data collected from the previous step includes all the review data along with posted date, rate, and information from the 64 airports. The reviews specifically contain the COVID-19 related keywords (Table 1), indicating how passengers thought about the airport during the COVID-19 pandemic in both positive and negative ways. Although the data collection has meaningful opinions from the 64 airports over 17 months, it is limited to manually investigate each individual review. Therefore, we use topic modeling techniques to systematically analyze the review data. Topic modeling employs machine learning and natural language processing to investigate the topics of a large collection of documents, and it was successfully adopted to analyze social media data to understand the public perceptions of certain topics (Hong and Davison, 2010).

#### 2.3.1. Preprocessing

The prerequisite step of topic modeling is preprocessing the text data into the right format. First, we converted all letters into lowercase. Then, we removed punctuation, emoji, and stop words. Although the reason for removing punctuation and emoji is clearly understandable, the list of stop words need to be tailored for our study specifically. First, it contains normal stop words such as pronouns, conjunctions, and any other grammatical words. The NLTK’s stop word list (Bird, 2006) was used to select such stop words. In addition, we also included the COVID-19 keywords (Table 1) and general airport related keywords (e.g., flight, airport, travel, etc.) in the list of stop words. This is mainly because these words are predominantly appeared in our review data due to our search method, which hinders to find the abstract topics. Lastly, we lemmatized our words in order to combine different formats of words into a singular one (Srinivasa-Desikan, 2018). After preprocessing, we then generated bigram and trigram words. In text mining, an n-gram is a group of adjoining n words from a given text or sentence. An n-gram with 1 word is a unigram; with 2 words, a bigram; with 3 words, a trigram. N-gram models are used to expand the context of the analysis and provide greater discrimination in the model (Manning and Schutze, 1999).

#### 2.3.2. Short text clustering

The most popular topic modeling algorithm is LDA, which is a generative probabilistic model to identify the abstract topic on text data (Blei et al., 2003). The two main assumptions of LDA are: 1) each document contains a distribution of multiple topics, and 2) each topic also consists of various words. Although LDA has been evidenced its performance of identifying abstract topics from various text data sources, it is challenging to cluster short text data (i.e., 1 - 2 sentence (30 - 50 words) documents) due to its sparse, high-dimensional, and large-volume characteristics (Qiang et al., 2020). In other words, short text data such as our data collection of Google Map review likely contain a single topic rather than a distribution of multiple topics, and this may reject the first assumption of LDA.

Therefore, we use the collapsed Gibbs Sampling based Dirichlet Multinomial Mixture model-based clustering (GSDMM) in our study. GSDMM was originally developed and introduced by Yin and Wang (2014). The core of the algorithm is that it assumes only one topic per document, which is suitable for short text clustering. Also, it is categorized as a soft clustering method, where we can calculate the probability (accuracy) of cluster assignments. To explain the algorithm, Yin and Wang employed an analogy of MovieGroup Process (Yin and Wang, 2014): A number of students are randomly assigned with K tables, and they are asked to re-choose a table in turn to group students with the same preferences of movie. Students have to follow two rules: 1) choose a table with more students (completeness), 2) choose a table whose students share similar interest (homogeneity). Asking this process multiple times, we expect that some tables grow larger, and others become smaller (or disappear). The experimental study of Google News and TweetSet datasets proved its clustering performance, which aims at the balance between completeness and homogeneity. The algorithmic details and experimental results can be found in Yin and Wang (2014).

#### 2.4. Spatial analysis

In addition to our topic modeling analysis, we spatially investigate the floor plans of the busiest airports in 2020 (i.e., ATL, CLT, DEN, DFW, LAX, ORD). The main objective is to investigate further on the selected airport regarding the traffic of passengers with respect to the comments of space issues from our topic modeling results.

We use Space Syntax, which is a set of theories to analyze spatial configurations and human activity patterns in buildings (Hillier and Hanson, 1984). The calculation is associated with how people perceive in large spaces. It is based on the concept that human attributes are directly associated with spatial configurations, which are used for understanding architectural and urban spaces to model human behaviors related to spatial elements (Karimi, 2018). Specifically, we study the agent run analysis of Space Syntax. It simulates the paths of the agents’ walk movements and returns the possible collapses of such paths, which is considered as the occupant density of the spaces.

We model the selected airports’ architectural geometry by AutoCAD software and simulated the agent run analysis using the software of DepthmapX (Turner, 2004), which is an open-source tool for Space Syntax theory. Although each airport consists of several terminals (or concourses), we select a single terminal (or concourse) from each airport to explain our simulation results effectively. The floor plans are traced as enclosed spaces for simplifying Space Syntax analysis, which assumes that passengers’ movements have only happened in the corridor and circulation in the terminal. This assumption is appropriate for analyzing terminal configuration because the corridor and circulation are the main spaces for passengers to move and wait. We use the number of passengers in 2020 for each airport as our simulation inputs.

### 3. Results

#### 3.1. Review data

The original dataset contains 103,428 reviews from the 64 airports over 17 months of data collection period (January, 2020 - May, 2021). After filtering the original review data with the COVID-19 related keywords in English, we are left with 7459 reviews to conduct our further analysis. The shortest and longest review text contain 1 and 675 words, respectively. The mean length of the review data is 33.5 words. As the 75% quartile of the review data is 41 words, the majority of these reviews are equal to or less than a sentence length. Therefore, we consider our analysis as a short text mining problem.

Table 2 presents the summary of our review data by Google Maps rating. More than half (71%) of the reviews are positive (1,885 EA for 4 star; 3413 EA for 5 star). This suggests that passengers generally had positive experience at the 64 airports regarding COVID-19 policy. The...
rest of neutral and negative reviews are 13% (965 EA for 3 star), 7% (535 EA for 2 star), and 9% (661 EA for 1 star) of the total data.

3.1.1. Policy impacts on the COVID-19 related reviews
In Fig. 5, we visualize and compare monthly airport passengers, the number of COVID-19 policy in place, and the COVID-19 related reviews. The number of airport passengers began to fall precipitously in response to the virus, developing health recommendations, and state policies as travel became much rarer. However, few COVID-19 related reviews were submitted at this point since expectations for public health standards were not yet well-established or understood. Shortly afterwards, many of the states in our dataset enacted stay-at-home orders, intuitively, the minimum number of passengers occurred when these orders were at their peak. Once states began to lift them and more people could travel again, the number of passengers began to increase, along with a corresponding increase in the COVID-19 related reviews. As expectations for health safety began to be better understood, more people began to discuss them in their airport reviews. Mask mandates were adopted around the US more gradually, but largely stayed in place until late in our data. Implementation of mask mandates is concurrent to an early increase in the number of passengers and reviews, followed by a long period of relative stability. Finally, travel quarantines were adopted by several states in our dataset, requiring visitors to self-quarantine after

Fig. 5. Policy impact timeline: the numbers of monthly passengers (up), adopted key policy (middle) and the COVID-19 related reviews (bottom).

Fig. 6. The top 30 frequently occurred words by rate information.
arriving. This disincentivized air travel and is aligned with a long period without any increase in passengers or reviews. After these restrictions began to be lifted, along with more widespread elimination of emergency status and mask mandates as states opened up, the number of passengers began to increase dramatically. The frequency of the COVID-19 related reviews varies over the duration of our dataset as driven by pandemic status and the policy landscape that developed in response.

3.1.2. Vocabulary of review data

We then obtained the vocabulary of our study by conducting the preprocessing steps defined in Subsection 2.3.1. The vocabulary is the list of unique words (7,202 words) by eliminating irrelevant words and adding N-gram words. Fig. 6 illustrates the top 30 frequently occurred words grouped by 1 - 5 star rate. For positive reviews, some words ('clean', 'easy', 'good', 'great', 'well', 'nice') were ranked at the most occurred words in both 4 and 5 star reviews. In particular, people acknowledged the word 'clean' for their positive experience at the airports. The words related to the airport functions were also frequently occurred (time: 'time', spaces: 'terminal', 'time', shops: 'restaurant', 'food', security: 'tsa', 'security'). It should be noted that these words are occurred in both positive and negative reviews. This leads us to further conduct the topic modeling analysis to investigate each function in detail.

Considering that each word can occur at 5 different rates, we calculate the weighted averaged of rate score of each word to understand the sentiment of each word by rate. We then select the most negative and positive words, whose scores were less than 2.5 and greater than 4.5, respectively (Fig. 7). Although most of the negative words contains innate negative meanings, we also found that they are related to some perceptive words of the airports ('shoulder', 'overcrowded', 'dirty', 'nowhere'). Similarly, the positive words contain some perceptive words of spaces ('spacious', 'plenty', 'quiet') of the airports. This suggests that individual ratings are often influenced by their experience of spatial issues in the airports. Spatial components of airport design can result in either very low, or very high, rated reviews. We will further analyze review comments with the perspective of space perceptions in
the next subsections.

### 3.2. Clustering results

Fig. 8 shows the clustering results of GSDMM on our text data. We investigate four results of GSDMM clustering assignment: 1) averaged accuracy of cluster assignment by iterating 100 times (Figure 8a), 2) standard deviation of the accuracy (Fig. 8b), 3) the maximum number of reviews assigned to clusters (Fig. 8c), and 4) the minimum numbers of reviews assigned to clusters (Fig. 8d). By increasing the number of clusters, the values of accuracy and standard deviation decrease and increase, respectively. Although it is important to maintain the high accuracy with lower standard deviation, we also need to increase the number of clusters to group the reviews with the associated topics for our interpretation. The maximum and minimum numbers of the assigned reviews for the clusters are stabilized around the cluster numbers of 4-5 clusters. After the cluster number of 5, it generates new clusters with a few minor reviews, which are not representative to indicate a certain topic. With the balance of four considerations, we select the cluster number of 4 to further analyze our text data.

Table 3 summarizes the result of GSDMM clustering. We assigned the topics of each cluster based on the word distribution of the four clusters.

In Fig. 9, the relative occurrence was sorted for each cluster (four different colored columns). For example, the word of ‘shop’ occurred the most in Cluster 1 with the ratio of 0.9 whereas it occurred the ratio of 0.1 in the three other clusters. Thus, we can infer that this word is the most important and representative word to understand the topic of Cluster 1.

On the other hand, Cluster 2 includes the words related to spatial topics (‘pack’, ‘full’, ‘crowd’, ‘seat’, ‘sit’, ‘bathroom’, ‘space’, ‘gate’, ‘area’), which is determined as the topic of ‘Space’.

Table 3 summarizes the result of GSDMM clustering. We assigned the topics of each cluster based on the word distribution of the four clusters.

In Fig. 9, Cluster 0 contains some positive words related to the operational aspect of the airport (‘organize’, ‘protocol’, ‘guideline’, ‘efficient’, ‘smooth’) as well as experiential words with other people (‘staff’, ‘follow’, ‘friendly’, ‘helpful’, ‘thank’, ‘pleasant’). Consider both categories of Cluster 0, we label it as the topic of ‘Staff’. More obviously, Cluster 1 contains the words related to shops in the airports (‘shop’, ‘store’, ‘restaurant’, ‘food’, ‘eat’, ‘drink’, ‘open’, ‘close’) relatively more than other clusters. Thus, Cluster 1 is considered as the topic of ‘Shop’.

In Fig. 9, Cluster 0 contains some positive words related to the operational aspect of the airport (‘organize’, ‘protocol’, ‘guideline’, ‘efficient’, ‘smooth’) as well as experiential words with other people (‘staff’, ‘follow’, ‘friendly’, ‘helpful’, ‘thank’, ‘pleasant’). Consider both categories of Cluster 0, we label it as the topic of ‘Staff’. More obviously, Cluster 1 contains the words related to shops in the airports (‘shop’, ‘store’, ‘restaurant’, ‘food’, ‘eat’, ‘drink’, ‘open’, ‘close’) relatively more than other clusters. Thus, Cluster 1 is considered as the topic of ‘Shop’.

On the other hand, Cluster 2 includes the words related to spatial topics (‘pack’, ‘full’, ‘crowd’, ‘seat’, ‘sit’, ‘bathroom’, ‘space’, ‘gate’, ‘area’), which is determined as the topic of ‘Space’. It should be noted that these...
spatial related words are rather negative meanings based on passengers’ incommodious special experiences. Lastly, we found other service related words ('car', 'parking', 'shuttle', 'bag', 'luggage', 'baggage claim', 'agent', 'ask', 'help', 'sign') and assigned the topic of 'Service' for Cluster 3. Cluster 0 (the topic of 'Staff') and 1 (the topic of 'Shop') contain 2465 and 2771 reviews, respectively, which account 70% of the total reviews. Cluster 2 (the topic of 'Space') and Cluster 3 (the topic of 'Service') take the rest of the reviews.

3.2.1. Detail investigations on clustering results

We then compare the clustering assignment with other characteristics of the airports (Fig. 10). There are more passengers (Fig. 10b) and
more COVID-19 confirmed cases of airport employees (Fig. 10d) in the airports of Cluster 2, while the actual sizes of the airports (inferred by the number of gates) were rather consistent (Fig. 10c). These observations of Cluster 2 suggest that the spatial complaints were not just driven by the architectural plans of the airports but also by the overcrowded number of passengers in the airports. There were no clear differences of political scores among 4 different clusters (Fig. 10e). The Google Maps rate on each cluster is also studied (Fig. 10a). The reviews in Cluster 0 have slightly more positive rates compared to other clusters. On the other hand, passengers reported their negative feedback on the topic of ‘Space’ in Cluster 2, which aligns well with the identified negative words in Figure 9.

Since it is critical to understand the sentiments of passengers in the airport, we will further investigate the relationship between the cluster assignment and rate in Fig. 11. The width of each bar indicates the number of reviews by cluster assignment (top) and rate (bottom). As shown in Fig. 11, the majority of reviews of Cluster 0 and 1 are positive reviews of 4 or 5 star rate. Especially, more than half of the reviews are ranked as 5 star rate in Cluster 0. This means that passengers valued their positive experiences of ‘Staff’ (Cluster 0) and ‘Shop’ (Cluster 1) by
indicating 4 or 5 star rate reviews. On the other hand, Cluster 2 and 3 contains relatively high proportion of negative reviews (1 or 2 star rate). Even worse the number of 1 star rate reviews is the most frequent in Cluster 2, which implies that passengers actively reported their negative impressions of spatial experiences during the COVID-19 pandemic. However, it should be explained that all 4 clusters have negative and positive reviews. Thus, it is limited to generalize the passenger experience of various aspects in the airports.

We also traced the number of reviews by cluster assignments and rate information with the monthly interval of our data collection period in Fig. 12. Prior to May, 2020, there were almost no reviews related to the COVID-19. This is because, only a very few passengers were aware of COVID-19 issues in January and February, 2020 and the volume of airport passengers were extremely low from March to April, 2020 due to the declaration of a national emergency in the US (March 13th, 2020) and associated policies. Beginning in May, 2020, more passengers started to report their airport experiences in Google Maps, and there are more reviews related to the topics of ‘Staff’ and ‘Shop’ compared to ‘Space’ and ‘Service’ during the COVID-19 pandemic period. As we discovered in Figure 11, Cluster 0 and 1 have more positive reviews, while Cluster 2 and 3 have a mixed ratio of 1 - 5 star rate reviews. Although the monthly number of reviews followed by the number passengers in Fig. 5, there was no clear temporal relationship of the number of reviews by cluster assignments and rate information.

### 3.2.2. Clustering results by airport

In Fig. 13, we analyze the results in detail with respect to the individual airport level. The left part of Figure 13 indicates the number of reviews (horizontal bar chart) and the average rate (line plot) of the COVID-19 related reviews on each airport. On average, one airport has 117 reviews with the COVID-19 keywords, which indicates that the concerns of the COVID-19 pandemic are spread out across the US. For major hub airports, ATL (Hartsfield-Jackson Atlanta International Airport) has the maximum (473 reviews) number of the COVID-19 related reviews. There are 26 airports with at least 100 reviews, which accounts for 76% of the total reviews. The average rates of the major hub airports are approximately around 4.0, while the lowest rate (3.0) is found from the reviews of CLT (Charlotte Douglas International Airport).

To investigate topical information of the reviews in detail, we visualize the clustering results in the right part of Fig. 13. The four individual horizontal bar charts indicate the number of reviews from Cluster 0 - 3 on each airport. Note that the scales of each bar charts are fixed to compare the quantities within the cluster. Similarly, the line plots show the average rates of the reviews of each cluster. In general, the rates of Cluster 0 and 1 are higher than the rates of Cluster 2 and 3, which confirms our finding in Subsection 3.2.1, i.e., people reported more negative experiences regarding space and service issues. Especially, the passengers of CLT airport reported the lowest rate score (2.0) of Cluster 2, which was the main reason to result in the minimum overall rate (3.0). Similarly, LAX (Los Angeles International Airport) and JFK (John
F. Kennedy International Airport) have relatively more reviews of Cluster 3. It should be noted that four cluster rate results of minor airports (upper part of Fig. 13) are hard to analyze in detail as the sample size is relatively small at such airports. In Section 4.1, we discuss the notable reviews of Cluster 2 and 3 to introduce newly arisen design and operation challenges of airports during/after the COVID-19 pandemic.

3.3. Detail investigations on spatial issues

We simulate the agent run analysis of Space Syntax (Hillier and Hanson, 1984) to investigate Space issues of the busiest airports. Each subplot of Fig. 14 indicates the results of the selected six airports which had the most passengers in 2020. We inset the floor plan on the left corner with the information of terminal and corridor. Then, we magnify the most crowded places on the bottom right side. The maximum density spots are visualized by red color, while blue color was used for the minimum density spots on each airport. It is important to note that this color legend is not applicable to compare the density of different airports.

All airports have similar patterns of passenger movements, i.e., maximum density at corridor spaces, which is main passenger circulation and located between shops, restaurants, and airport facilities. In addition, they all follow the airport planning and design guidebook (Board et al., 2010) by designing clear circulation width by 20-feet wide for single-loaded concourses and 30-feet wide for double-loaded concourses. However, in CLT, the corridor is relatively narrower than in other airports. This spatial characteristic may suggest why CLT has more reviews in Cluster 2 in our text mining results. Our results indicate that agent movement density in LAX is evenly distributed around terminal spaces. This is mainly because the terminal layout that is not formed a geometric boundary but an irregular configuration. ATL and DEN (Denver International Airport) have the longest passenger walking distances. The graph illustrates high-density area is only located at the center corridor, which is used for moving walkways. The terminals are very linear and straight, where they assign the various programs appropriately to avoid the high-density spots. DFW (Dallas Fort Worth International Airport) is a curved shape, but it is very similar to the other two linear shapes.

4. Discussion

4.1. Recommendations from individual reviews

The proposed text mining based approach provides interesting insights to understand public perceptions with respect to FM in airports. In this regard, we cite individual reviews found in our data collection as evidence recommending future points of airports. The reviews of Cluster 2 and 3 are relatively negative, and we investigate their implications further in our discussion. In summary, there are four areas: 1) spatial planning and design, 2) gate assignment and operation, 3) policy, and 4) building maintenance. We detail each category with actual passengers’ reviews as follows.

4.1.1. Spatial planning and design

In comments regarding spatial planning and design, individuals who indicated their negative reviews mentioned that the airports were overcrowded and not spacious enough for the passengers. These suggest possible architectural design features of waiting areas in buildings (Choi, 2021; Ugal et al., 2021). With COVID-19 restrictions, passengers had not been allowed to eat food in some airplanes, which made people instead choose to eat at the airport before boarding. Accordingly, it required more sitting areas in the airport. Also, some major airports were under renovation or construction, making parts of the waiting areas unavailable and leading to detours of passengers’ foot traffic.

“Impossible to have any social distancing. Crowded, narrow hallways and doorways” “I feel like the hallways in the terminals should have been built wider” “The airport was over crowded and dirty. Not enough sitting space to support social distancing” “Current renovations made makeshift walkways very narrow for passing other travelers to get to our next gate” “I understand there is construction going on, but there must be a way to better regulate foot traffic for social distancing”

4.1.2. Gate assignment and operation

In comments regarding gate assignment and operation, passengers complained that some gate areas were more crowded than others, which might be improved by occupancy-based operation. This suggests that the airport management team can more intelligently assign the gates to avoid overcrowded occupancy traffic in terminal side. Although it should be noted that airport gate management involves more complex strategies (Bouras et al., 2014), such as available taxiing areas of the airplanes, the volume of occupant traffic should be considered in the gate assignment and operation, particularly during health crises such as the COVID-19 pandemic.

“there were many gates crammed in a small area, all boarding at the same time with no way to distance” “Social distancing at the gate is near impossible, unless you are willing to wait back in the hallway” “lots of confined areas, multiple gates from the same relatively small area (12 gates in this particular spot), and there are very limited seating/waiting areas”

4.1.3. Airport service policy

In comments regarding airport policy, passengers mentioned that overcrowding also occurred due to the lack of particular policies in the airports, such as rules for elevator usage, internal transportation systems, and signage. Although the main modes of foot traffic are escalators and moving walks, elevators are still important, especially for people who have physical challenges. Thus, restrictive rules for elevator usage should be designed and maintained. In addition, large hub airports have internal transportation systems (e.g., shuttles, trains) connecting multiple terminals. Some passengers commented that there were no clear instructions on such transportation systems to operate (e.g. maximum occupancy). It is recommended to develop and implement guidelines for internal transportation systems. Lastly, passengers reported that the signage was not informative enough to understand COVID-19 policy, including regulations regarding social distancing in sitting areas. For better communication, it is required to revisit current airport signage to improve their clarity and interactive aspects.

“Escalators don’t work to get to different levels forcing packed elevators not allowing social distancing guidelines required by the government” “No limiting of passengers on the trams” “Misleading signs taking you to nowhere” “The signs are very last minute and not helpful or intuitive”

4.1.4. Building maintenance

In comments regarding cleaning, several passengers reported that there were cleaning issues regarding hand sanitizers and soaps. Although such complaints have existed even before the COVID-19 pandemic, passengers consider this problem more seriously by linking it with the virus transmissions. Thus, it is required to improve the protocols of airport cleaning and clearly explain that to the passengers. The actual review comments also suggested implementing touch-free technologies for hand sanitizers and soaps.

“Spent night (6 h layover) ... no covid cleaning” “Poor COVID response. Few hand sanitizer stations and many of them were empty” “We have to touch kiosks that everyone else touches, remove masks
and glasses and they provide no hand sanitizer. All hand sanitizer dispensers are empty.” “The restrooms do not have touch free fixtures”

In comments regarding building environmental controls, some passengers reported that they felt hot and humid by having their mandatory masks on. Such comments suggested that optimal temperatures in the airports should be redefined to provide thermal comfort and healthy indoor environmental quality considering occupant comfort carefully (Jazizadeh et al., 2014; Park and Nagy, 2018; 2020).

“You were required to wear your mask at all time for personal safety but they have the temperature about 78 ° throughout the airport which is making everybody sweat extremely bad. If you want people to wear their masks then you need to make it a comfortable temperature so that they don’t mind wearing the mask” “Better than most, but they pump air in from directly outside, so sitting in the row of rocking chairs is unpleasant as exhaust is blown in your face” “It is grossly overcrowded especially during a pandemic, the climate control was a joke”

4.2. Implications for building design and operation

The COVID-19 pandemic imposed us to reconsider our building design and safety features in built environment (Awada et al., 2021). Multiple stakeholders involved in the building project lifecycle (i.e., design, construction and operation) have made efforts toward the healthier built environment. However, there are still necessities for effective and practical strategies to prevent virus transmission and improve the compliance of the health-related policies (e.g., NPI), especially in large facilities such as airports (Choi, 2021; Serrano and Kazda, 2020). Our study provides a closer investigation into the perceptions of airport passengers with considerations of multifaceted human-building interactions.

Regarding the prevention of occupants’ disease transmission, it is critical to understand spatial dynamics, building operational factors, and occupant behavior. Especially, space configurations critically influence on human building interactions. Thus, we can employ the knowledge of spatial interactions of occupants to the decision-making process regarding disease transmission and social-distancing measures in buildings (Dietz et al., 2020). For example, the comments of Cluster 2, indicated that the distance between each gate is too short for passengers to keep social distancing. In other words, passengers felt overcrowded in hallways, sitting spaces, and boarding gates. Despite the gates and concourses were planned to follow the airport design and planning guidelines, it is not desirable for people to support the social distance in an airport terminal during pandemic.

This observation further requires developing new guidelines being drawn up to determine the shape of future terminals and develop terminal planning design for the post COVID-19 era. The Airports Council International (ACI) and the International Air Transport Association (IATA) have published guidance documents to manage COVID-19. The reports describe a set of typical measures to be implemented at major international airports; staff management, social distance policy, passenger management, facility planning, and security screening. This assists airports to mitigate the impact of COVID-19 and reduce in queuing capacity which means increasing queuing areas in circulation zone. We anticipate that our results should be suggested for such guidelines to incorporate public perceptions of COVID-19 policy.

4.3. Implications for public policy

Few issues in public life and public health policy have been as contentious as COVID-19 policy. Transportation hubs were identified early on as places where transmission of the virus could occur and spread the illness across vast stretches of land. Based on comments from our data, we have identified several possible changes in policy: more spacious areas to allow for social distancing by passengers, shifting the way gates and departures/arrival slots are allocated so fewer passengers are in crowded gate areas, better posting of policies for use of elevators, escalators, walkways, and more attention to maintenance of anti-COVID facilities and measures. Of course, few such policies are self-enforcing, so implementation and enforcement of any such policies is vital.

These types of policies are classified as regulative (Lowi, 1972) and are generally created by engineers or others with specialized expertise. To have such changes will require attitudinal shifts in education of the experts in charge (in school) which need to be updated as conditions shift. Legislators are understandably unlikely to want to legislate the size and layout of airports (and other transportation hubs), the ways aircraft are queued, and gates allotted. These standards are appropriately assigned to lower-level experts to promulgate. Review of these policies should occur at regular intervals, however, and, as we have shown, members of the general public should be able to weigh in on how well the policies work from the consumer perspective.

The methodology we develop in this study demonstrates the utility of social media reviews for understanding user-based experience of different facilities. This model may be useful to practitioners, helping them to integrate input from the public in their work. Facility managers can use this method to access public perceptions, making them better able to identify areas of weakness, anticipate and respond to issues, and assess user compliance to implemented policies. If, or when, future crises occur, putting strain on system operations, these abilities can help managers adapt their facilities to the situation as it evolves. With more flexible, data-driven responses to health crises, managers may be able to optimize facility processes more accurately and implement policies more effectively in order to keep operations running, even if at lowered capacity, without sacrificing user health. Furthermore, understanding how users’ perceptions unfold in each unique space can help policymakers and managers develop more effective policy instruments promoting safe behavior in public spaces. This can help promote positive user perceptions of transportation hubs, maintain healthier public spaces, and increase resiliency of the transportation sector to future health crises. In future, facility managers should implement systems using this method of user feedback monitoring to make their spaces more amenable to being adapted to novel circumstances, making them more resilient in the face of crisis.

Policy nudges are a powerful tool for encouraging acceptable public behavior and are extremely useful in situations where direct regulation or enforcement are not present, such as with public health safety during the COVID-19 pandemic in many places in the US. However, the impact of building design and architecture on individuals’ responses to behavioral policies has rarely been studied. Our results show the importance of architectural space in how individuals perceive and respond to COVID-19 policy, indicating the importance of design in the build environment on how individuals act. Building design defines the choice architecture individuals are presented with, either constraining or facilitating how policy nudges work. Identical policies in different architectural spaces can lead to very different outcomes in terms of how individuals choose to act. In this way, building architecture moderates choice architecture. Future researchers investigating the impact of behavioral cues as policy initiatives have to take these factors into account. Likewise, policymakers must consider the unique properties of the designed space their policy will be implemented in if they want to achieve effective outcomes. Failure to do so risks the efficacy of the policy and may result in counterproductive results. Policies must be tailored to the space they are in when being implemented.

4.4. Limitations

Although the results were based on one of the largest datasets of public opinions in airports, the generalizability of our findings remains limited, which is a common limitation in any data-driven studies using
building data (Miller et al., 2018; Park et al., 2019c). Use of other social media resources (e.g., Twitter, Instagram) to gather and analyze for understanding the human perceptions of COVID-19 policy may be useful to validate these findings. Moreover, there are other population groups who barely express their opinions via social media and are thus excluded from our sample. Therefore, the results should be carefully interpreted, as additional positive and/or negative perceptions regarding COVID-19 policy may exist. Even with this limitation, it should be reiterated that our dataset (103,428 reviews) is unique and suitable for the objective of this article because Google Maps specifically aims to collect the reviews of certain locations (e.g., airports, restaurants, parks). For future study, we plan to compare different social media options to understand the human perceptions of different transportation hubs as well.

For the spatial analysis of each airport terminal, we used Space Syntax. Since the airport terminal is a directional space, Space Syntax, which is based on the random patterns of movements, has limitations to predict the airport passenger movements accurately. Furthermore, the movements are dynamic and are affected by the spatial environment. Thus, our simulation approach has challenges to demonstrate how to respond to more detailed spatial configurations (Esposito et al., 2020).

5. Conclusion

Airports are key vectors in the spread of contagious diseases, where they present health risks to visitors. While policies requiring NPIs to limit the spread of the coronavirus have been implemented, safety levels in different airports are highly variable. In this article, we investigate the public perception of COVID-19 policy at airports by employing text mining techniques to identify how user perceptions differ, key issues related to the pandemic, and potential drivers in heterogeneous safety public perception of COVID-19 policy at airports by employing text mining techniques to identify how user perceptions differ, key issues related to the pandemic, and potential drivers in heterogeneous safety levels among airports. We collected 103,428 unique Google Maps reviews from 64 hub airports in the US over a 17-month period (January, 2020 - May, 2021), filtering our raw dataset by the COVID-19 related key-

words, resulting in a final set of 7459 reviews. Using GSDMM clustering, we identify four different topics of interest that shed light on how COVID-19 has affected facility management: Staff (Cluster 0), Shop (Cluster 1), Space (Cluster 2) and Service (Cluster 3). In general, airport passengers express having positive experiences with Staff and Shop, but neutral or negative experiences with Service and Space. This demonstrates the extent of concerns airport visitors have identified with Space and Service issues because of the behavioral changes demanded by the pandemic. Furthermore, it shows the impact of building design and operation on visitor compliance to facility health policies, indicating how buildings modify user behavior and health safety. Based on the negative reviews in Clusters 2 and 3, we discuss potential approach for facility managers to improve 1) spatial design and planning, 2) gate assignment and operation, 3) airport service policy, and 4) building maintenance, in order to adapt airports to COVID-19 conditions and prepare them in case of future health crises.

The proposed text mining-based framework on social media data provides an innovative approach to facilitate human-centric facility management and offers a viable strategy to monitor and adapt health policy implementation to large transportation infrastructures (e.g., airports). Eventually, this can help lower non-compliance to facility health policies and mitigate the discrepancy between health policy goals and perceptions of their efficacy, shaping healthy and safe urban infrastructures for the future.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A

Table 4

| Code | Location      | Hub     | Terminal Type | Nr. Gate | Nr. Covid19 cases (TSA) | No. Passenger (2020) |
|------|---------------|---------|---------------|----------|-------------------------|----------------------|
| ANC  | Anchorage, AK | M       | Linear, Pier  | 8        | 19                      | 1,157,378            |
| ATL  | Atlanta, GA   | L       | Satellite     | 195      | 190                     | 20,559,866           |
| AUS  | Austin, TX    | L       | Pier          | 24       | 56                      | 6,288,519            |
| BDL  | Windsor Locks, CT | M   | Pier          | 37       | 34                      | 1,150,033            |
| BNA  | Nashville, TN | L       | Pier          | 42       | 56                      | 4,016,808            |
| BOI  | Boise, ID     | M       | Pier          | 23       | 29                      | 992,342              |
| BOS  | Boston, MA    | L       | Linear, Pier  | 102      | 168                     | 6,035,452            |
| BUR  | Burbank, CA   | M       | Linear        | 23       | 32                      | 1,056,838            |
| BWI  | Baltimore, MD | L       | Pier          | 77       | 86                      | 5,452,563            |
| CHS  | Charleston, SC| M       | Pier          | 15       | 12                      | 945,672              |
| CLE  | Cleveland, OH | M       | Pier          | 16       | 42                      | 1,994,909            |
| CLT  | Charlotte, NC | L       | Pier          | 115      | 95                      | 12,952,869           |
| CMH  | Columbus, OH  | M       | Pier          | 34       | 35                      | 1,577,596            |
| CVG  | Covington, KY | M       | Satellite     | 51       | 33                      | 1,746,546            |
| DAL  | Dallas, TX    | M       | Linear        | 20       | 54                      | 3,669,930            |
| DCA  | Arlington, VA | M       | Pier          | 59       | 83                      | 3,573,489            |
| DEN  | Denver, CO    | L       | Satellite     | 146      | 125                     | 16,234,349           |
| DFW  | Dallas-Fort Worth, TX | L | Pier | 182 | 240 | 18,593,421 |
| DTW  | Detroit, MI   | L       | Satellite     | 129      | 101                     | 6,822,324            |
| EWR  | Newark, NJ    | L       | Pier          | 121      | 323                     | 7,986,224            |
| FLL  | Ft. Lauderdale, FL | L | Pier | 66 | 276 | 8,022,354 |
| HNL  | Honolulu, HI  | M       | Pier          | 48       | 22                      | 3,126,391            |
| HOU  | Houston, TX   | M       | Linear        | 30       | 45                      | 3,127,178            |
| IAD  | Chantilly, VA | L       | Satellite     | 113      | 83                      | 3,862,658            |
| IAH  | Houston, TX   | L       | Pier          | 130      | 171                     | 8,682,558            |
| IND  | Indianapolis, IN | M  | Pier          | 33       | 35                      | 1,999,865            |
| JAX  | Jacksonville, FL | M   | Linear        | 20       | 31                      | 1,369,681            |
| JFK  | New York, NY  | L       | Pier, Satellite | 131 | 395 | 8,269,819 |
| LAS  | Las Vegas, NV | L       | Pier, Satellite | 92 | 193 | 10,594,701 |
| LAX  | Los Angeles, CA | L   | Pier          | 132      | 439                     | 14,059,754           |
| LGA  | New York, NY  | L       | Pier          | 86       | 150                     | 4,147,116            |
| MCI  | Kansas City, MO | M   | Satellite     | 39       | 29                      | 2,169,853            |
| MCO  | Orlando, FL   | L       | Satellite     | 129      | 256                     | 10,467,728           |
| MDW  | Chicago, IL   | L       | Pier          | 43       | 79                      | 4,237,025            |
| MEM  | Memphis, TN   | M       | Pier, Linear  | 31       | 23                      | 1,019,829            |
| MIA  | Miami, FL     | L       | Pier          | 131      | 373                     | 8,786,007            |
| MKE  | Milwaukee, WI | M       | Pier          | 38       | 30                      | 1,263,385            |
| MSP  | Minneapolis, MN | L         | Pier          | 131      | 88                      | 7,069,720            |
| MSY  | New Orleans, LA | M   | Linear        | 35       | 57                      | 2,633,259            |
| OAK  | Oakland, CA   | M       | Pier          | 32       | 26                      | 2,271,706            |
| OGG  | Kahului, HI   | M       | Linear        | 39       | 3                       | 1,135,141            |
| OKC  | Oklahoma City, OK | M | Linear | 16 | 20 | 936,241 |

(continued on next page)
Table 4 (continued)

| Code | Location        | Hub     | Terminal Type | Nr. Gate | Nr. Covid19 cases (TSA) | No. Passenger (2020) |
|------|-----------------|---------|---------------|----------|-------------------------|----------------------|
| OMA  | Omaha, NE       | M       | Satellite     | 20       | 26                      | 1,037,753            |
| ONT  | Ontario, CA     | M       | Satellite     | 26       | 29                      | 1,237,946            |
| ORD  | Chicago, IL     | L       | Pier          | 191      | 292                     | 1,606,015            |
| PBI  | West Palm Beach, FL | M       | Pier          | 34       | 37                      | 1,519,766            |
| PDX  | Portland, OR    | M       | Pier          | 51       | 28                      | 3,455,877            |
| PHL  | Philadelphia, PA | L       | Pier          | 126      | 161                     | 5,753,156            |
| PA   | Phoenix, AZ     | L       | Pier          | 106      | 180                     | 10,531,436           |
| PIT  | Pittsburgh, PA  | M       | Satellite     | 75       | 38                      | 1,749,592            |
| RDU  | Raleigh, NC     | M       | Linear        | 45       | 31                      | 2,327,816            |
| RNO  | Reno, NV        | M       | Pier          | 23       | 11                      | 977,838              |
| RSW  | Ft. Myers, FL   | M       | Linear        | 28       | 65                      | 2,947,139            |
| SAN  | San Diego, CA   | L       | Pier          | 51       | 84                      | 4,637,856            |
| SAT  | San Antonio, TX | M       | Pier          | 24       | 51                      | 1,920,042            |
| SEA  | Seattle, WA     | L       | Pier          | 103      | 110                     | 9,462,173            |
| SFO  | San Francisco, CA | L       | Pier          | 115      | 57                      | 7,745,057            |
| SJC  | San Jose, CA    | M       | Linear        | 17       | 31                      | 2,283,186            |
| SJC  | San Jose, CA    | M       | Linear        | 38       | 39                      | 2,356,948            |
| SLC  | Salt Lake City, UT | L     | Pier          | 71       | 91                      | 5,981,032            |
| SMF  | Sacramento, CA  | M       | Linear        | 40       | 38                      | 2,710,342            |
| SNA  | Santa Ana, CA   | M       | Linear        | 20       | 35                      | 1,824,836            |
| STL  | St Louis, MO    | M       | Pier          | 82       | 53                      | 3,035,469            |
| TPA  | Tampa, FL       | L       | Satellite     | 59       | 76                      | 4,966,775            |

References

Administration, F. A. (2018). Advisory circular (ac 150 / 5360-13a). Airport Terminal Planning.
Ali, S. T., Wang, L., Lau, E. H., Xu, X.-K., Du, Z., Wu, Y., … Cowling, B. J. (2020). Serial interval of sars-cov-2 was shortened over time by nonpharmaceutical interventions. Science (New York, N.Y.), 369(6507), 1106-1109.
Awada, M., Becerik-gerber, B., White, E., Hosque, S., O’neill, Z., Pedrielli, G., … Wu, T. (2021). Occupant health in buildings: Impact of the covid-19 pandemic on the opinions of building professionals and implications on research. Building and Environment, 108440.
Banholzer, N., Van Weenen, E., Kratzwald, B., Seeliger, A., Tschernutter, D., Bottrighi, P., … Ali, S. T., Wang, L., Lau, E. H., Xu, X.-K., Du, Z., Wu, Y., … Wu, T. (2020). 2019 novel coronavirus (covid-19) pandemic: built environment considerations to reduce transmission. Mojave, 52(2), e0245-20.
Dutta, S., Gunay, H. B., & Bucking, S. (2021). Benchmarking operational performance of buildings by text mining tenant surveys. Science and Technology for the Built Environment, 27(6), 741–755.
Ebeid, J. J., & Kolsi, B. (2020). Public risk perception and emotion on twitter during the covid-19 pandemic. Applied Network Science, 5(1), 1-32.
Edelson, P. J., & Phypers, M. (2011). Tbias transmission on public transportation: a review of published studies and recommendations for contact tracing. Travel Medicine and Infectious Disease, 9(1), 27-31.
Elachola, H., Ebrahim, S. H., & Gozzer, E. (2020). Covid-19: Face mask use prevalence in international airports in asia, europe and the americas, march 2020. Travel Medicine and Infectious Disease, 35, 101637.
Espoto, D., Santoro, S., & Camardia, D. (2020). Agent-based analysis of urban spaces using space syntax and spatial cognition approaches: A case study in bari, italy. Sustainability, 12(11), https://doi.org/10.3390/su12114625
Frontzek, M., Schiavon, S., Goins, J., Aresen, E., Zhang, H., & Wargocki, P. (2012). Quantitative relationship of occupant satisfaction and satisfaction with indoor environmental quality and building design. Indoor air, 22(2), 119-131.
Gadain, S. K., Goodman, S. W., & Pepinsky, T. B. (2021). Partisanship, health behavior, and policy attitudes in the early stages of the covid-19 pandemic. PloS one, 16(4), e0249096.
Gaskin, D. J., Zare, H., & Delmarter, B. A. (2021). Geographic disparities in covid-19 infections and deaths: The role of transportation. Travel policy, 102, 35-46.
Goins, J., & Moezzi, M. (2013). Linking occupant complaints to building performance. Building Research & Information, 41(3), 361-376.
Gude, V. G., & Muire, P. J. (2021). Preparing for outbreaks–implications for resilient v wastewater operations and services. Sustainable Cities and Society, 64, 102558.
Gunay, H. B., O’Brien, W., Beausoleil-morrison, I., & Bursil, J. (2016). Development and implementation of a thermo artificial learning algorithm. Science and Technology for the Built Environment, 24(1), 43-56.
Hillier, B., & Hanson, J. (1984). The social logic of space. Cambridge University Press.
Hillier, L., & Davidson, B. D. (2010). Empirical study of topic modeling in twitter. Proceedings of the first workshop on social media analytics (pp. 80-88).
Hornik, R., Kikut, A., Jesch, E., Woko, C., Siegel, L., & Kim, K. (2021). Association of covid-19 misinformation with face mask wearing and social distancing in a nationally representative us sample. Health communication, 36(1), 6-14.
Ikonen, N., Savolainen-kopra, C., Enstone, J. E., Kulmala, I., Pasanen, P., Salmela, A., … Ruutu, P. (2018). Deposition of viral respiratory pathogens on frequently touched surfaces at airports. BMC infectious diseases, 18(1), 1-7.
Jazizadeh, F., Ghahramani, A., Becerik-Gerber, B., Keykhaey, T., & Orosz, M. (2014). User-led decentralized thermal comfort driven hvac operations for improved efficiency in office buildings. Energy and Buildings, 70, 398-410.
Keil, R., & Ali, S. H. (2011). The urban political pathology of emerging infectious disease in the age of the global city. Urbanism: Cities & Policymaking in the Global Age, 123-145.
Kithore, N., Mitchell, R., Lash, T. L., Reed, C., Danon, L., Sigmundsdottir, G., & Vigfusson, Y. (2020). Flying, phones and flu: Anonymized call records suggest that keflavik international airport introduced pandemic h1n1 into iceland in 2009. Influenza and other respiratory viruses, 14(1), 37-45.
Kurian, S. J., Alvi, M. A., Ting, H. H., Storlie, C., Wilson, P. M., Shah, N. D., Liu, H., … Keil, R., & Ali, S. H. (2011). The urban political pathology of emerging infectious disease in the age of the global city. Urbanism: Cities & Policymaking in the Global Age, 123-145.
Lavy, S., Garcia, J. A., & Dixit, M. K. (2010). Establishment of kpi for facility performance measurement: review of literature. Facilities.
Leaman, A., & Bordass, B. (2007). Are users more tolerant of greenbuildings? Building Research & Information, 35(6), 662-673.
Lee, K., & Yu, C. (2018). Assessment of airport service quality: A complementary approach to measure perceived service quality based on google reviews. Journal of Air Transport Management, 71, 28–44.
Lowi, T. J. (1972). Four systems of policy, politics, and choice. Public Administration Review, 32(4), 298–310.
Mangan, G., & De Marco, A. (2014). The role of maintenance and facility management in logistics: A literature review. Facilities.
Manning, C., & Schutze, H. (1999). Foundations of statistical natural language processing. MIT press.
Megahed, N. A., & Ghoneim, E. M. (2020). Antivirus-built environment: Lessons learned from covid-19 pandemic. Sustainable cities and society, 61, 102550.
Miller, C., Nagy, Z., & Schlaeter, A. (2018). A review of unsupervised statistical learning and visual analytics techniques applied to performance analysis of non-residential buildings. Renewable and Sustainable Energy Reviews, 81, 1365-1377.
Mo, Y., Zhao, D., Du, J., Sysal, M., Aziz, A., & Li, H. (2020). Automated staff assignment for building maintenance using natural language processing. Automation in Construction, 119, 103150.

Moezzi, M., & Goins, J. (2011). Text mining for occupant perspectives on the physical workplace. Building Research & Information, 39(2), 169–182.

Nakamura, H., & Managi, S. (2020). Airport risk of importation and exportation of the covid-19 pandemic. Transport policy, 96, 40–47.

O’Brien, W., Wagner, A., Schweiker, M., Mahdavi, A., Day, J., Kjærgaard, M. B., …, Yan, D., et al. (2020). Introducing iea ebc annex 79: Key challenges and opportunities in the field of occupant-centric building design and operation. Building and Environment, 178, 106738.

Park, J. Y., Dougherty, T., Fritz, H., & Nagy, Z. (2019a). Lightlearn: An adaptive and occupant centered controller for lighting based on reinforcement learning. Building and Environment, 147, 397–414.

Park, J. Y., & Nagy, Z. (2018). Comprehensive analysis of the relationship between thermal comfort and building control research-a data-driven literature review. Renewable and Sustainable Energy Reviews, 82, 2664–2679.

Park, J. Y., & Nagy, Z. (2020). Hvacelearn: A reinforcement learning based occupant-centric control for thermostat set-points. Proceedings of the eleventh ACM international conference on future energy systems (pp. 434–437).

Rovetta, A., Bhagavathula, A. S., et al. (2020). Global infodemiology of covid-19: analysis of google web searches and instagram hashtags. Journal of Medical Internet Research, 22(8), e20673.

Ridhwan, K. M., & Hargreaves, C. A. (2021). Leveraging twitter data to understand public sentiment for the covid-19 outbreak in singapore. International Journal of Information Management Data Insights, 100021.

Rovetta, A., Bhagavathula, A. S., et al. (2020). Global infodemiology of covid-19: analysis of google web searches and instagram hashtags. Journal of Medical Internet Research, 22(8), e20673.

Roy, S., & Ghosh, P. (2020). Factors affecting covid-19 infected and death rates inform lockdown-related policymaking. PloS one, 15(10), e0241165.

Saleh, S. N., Lehmann, C. U., McDonald, S. A., Basit, M. A., & Medford, R. J. (2021). Understanding public perception of coronavirus disease 2019 (covid-19) social distancing on twitter. Infection Control & Hospital Epidemiology, 42(2), 131–138.

Serrano, F., & Kazda, A. (2020). The future of airport post covid-19. Journal of Air Transport Management, 89, 101900.

Serrano, F., & Kazda, A. (2020). The future of airport post covid-19. Journal of Air Transport Management, 89, 101900.

Srinivasa-Desikan, B. (2018). Natural language processing and computational linguistics: A practical guide to text analysis with python, gensim, spaCy, and keras. Packt Publishing Ltd.

Thaler, R. H., & Sunstein, C. R. (2018). Nudge: Improving decisions about health, wealth, and happiness. HeinOnline.

Tucker, M., & Pitt, M. (2009). Customer performance measurement in facilities management: a strategic approach. International journal of productivity and performance management.

Turner, A. (2004). Depthmap 4: A researcher’s handbook.

Ugail, H., Aggarwal, R., Iglesias, A., Howard, N., Campuzano, A., Suárez, P., Maqsood, M., Aalid, F., Mehmood, I., Gleighorn, S., et al. (2021). Social distancing enhanced automated optimal design of physical spaces in the wake of the covid-19 pandemic. Sustainable Cities and Society, 68, 102791.

Witten, I. H., & Frank, E. (2002). Data mining: practical machine learning tools and techniques with java implementations. ACM Sigmod Record, 31(1), 76–77.

Yin, J., & Wang, J. (2014). A dirichlet multinomial mixture model-based approach for short text clustering. Proceedings of the 20th ACM SIGKDD international conference on knowledge discovery and data mining (pp. 233–242).