Computers, Environment and Urban Systems 130 (2026) 102507

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/ceus

Evaluating the digital divide in public transport services: A natural language processing analysis of user experiences in Birmingham ☆

Lu Huafeng [a,*], Lei Zhang [b,c,d,e], Silvia Gullino [a], Miguel Hincapié Triviño [a]

[a] Department of Architecture and Built Environment, School of Architecture, Built Environment, Computing and Engineering, Birmingham City University, Birmingham B4 7BD, UK
[b] School of Psychology, University of Birmingham, Birmingham B15 2TT, UK
[c] Centre for Human Brain Health, University of Birmingham, Birmingham B15 2TT, UK
[d] Institute for Mental Health, University of Birmingham, Birmingham B15 2TT, UK
[e] Centre for Developmental Sciences, University of Birmingham, Birmingham B15 2TT, UK

ARTICLE INFO

Keywords:
Digital divide
Public transport
Smart City
Inclusive planning
Natural language processing
Artificial intelligence

ABSTRACT

Digitalisation has transformed public transport services worldwide, yet disparities in access, skills, and outcomes continue to reinforce a digital divide in how citizens engage with these services. While previous studies have examined the digital divide in internet use and smart technologies, few have explored how these inequalities manifest in everyday interactions with digital public transport platforms, extending to citizen participation in government-provided digital services. To this end, we investigated user experiences of public transport applications in Birmingham, whose digital inclusion is disproportionate to its status as a major urban and economic hub in the UK, by analysing 4275 user reviews on public transport services collected from the Apple App Store and Google Play Store between 2016 and 2025. Using a mixed natural language processing (NLP) approach combining topic modelling and sentiment analysis, we identified key themes and emotional tones in user feedback. Results revealed three dominant topics: App Performance, Usability, and Service Satisfaction, each conceptually aligned with the access, ability, and motivation dimensions of the digital divide theory. Sentiment analysis showed generally negative evaluations, particularly around technical access barriers and satisfaction issues. By integrating user-generated data with digital divide theory, these results provide novel empirical evidence of how inequalities in digital engagement persist within digital twin planning and smart city transport systems and highlight the importance of inclusive digital design for equitable and sustainable mobility access.

1. Introduction

The digital transformation of public transport systems has become a cornerstone of "smart city" development (Edwards & Mackett, 1996; Shekarappa et al., 2023; Wilson et al., 2019), promising greater efficiency, user convenience, and sustainability. Mobile applications for journey planning, ticketing, and service updates are now integral to how citizens interact with urban mobility networks (Frederico et al., 2021). These technologies are often promoted as key enablers of inclusive and sustainable urban governance; however, these benefits are not distributed equally (Durand et al., 2022). Despite widespread smartphone adoption, digital participation in public transport remains uneven across regions and demographic groups. In the UK, Birmingham exemplifies this imbalance: as one of the country's largest and most diverse cities, it continues to rank among the most digitally excluded urban areas (Allmann & Radu, 2023; Comber et al., 2022; Faith et al., 2022). This discrepancy highlights a broader paradox of digital urbanism, while digitalisation is often celebrated as a tool for inclusion and empowerment, it can simultaneously reproduce or deepen existing social and spatial inequalities (Helsper, 2021).

This persistent unevenness reflects the digital divide, a multidimensional concept that captures disparities across three interrelated levels (Hargittai, 2003; Van Deursen & Helsper, 2015; Van Dijk, 2006), to reflect not only in access to digital technologies but also in the skills required to use them and the outcomes people derive from their use. The first-level digital divide, or access divide, concerns the physical and financial ability to obtain and use digital technologies, such as owning a smartphone, maintaining stable data access, or having reliable network

☆ This article is part of a Special issue entitled: 'Digital Twins and AI for Cities' published in Computers, Environment and Urban Systems.

* Corresponding author.
E-mail address: huafeng.lu@mail.bcu.ac.uk (L. Huafeng).

https://doi.org/10.1016/j.compenvurbsys.2026.102507
Received 28 November 2025; Received in revised form 13 July 2026; Accepted 10 August 2026
Available online 19 August 2026

coverage for transport apps. The second level, skills divide, reflects disparities in the knowledge and ability required to navigate digital platforms effectively, including app literacy, troubleshooting capacity, and the confidence to use digital tools for everyday mobility. The third-level digital divide, or outcome divide or motivation divide, describes differences in the benefits individuals obtain from digital engagement, such as improved travel efficiency, satisfaction, or trust in digital governance systems. Together, these three layers illustrate that digital exclusion is not merely a question of technology availability, but of how citizens are enabled, or sometimes disabled, in translating digital opportunities into meaningful digital engagement in public activities (Helsper, 2021; Nguyen & Hargittai, 2023; Nie et al., 2024; Nurminen et al., 2024).

In the context of public transport services, the digital divide has noticeable consequences. As booking, ticketing, and service updates increasingly shift to online or digital platforms, digital literacy becomes a prerequisite for everyday mobility (e.g., Durand et al., 2022). Studies have shown that individuals with limited digital skills or access are less able to benefit from innovations such as contactless payments or multi-function journey planning apps, effectively creating new barriers to mobility (Faith et al., 2022). These challenges undermine the social goals of smart cities, where digitalisation is intended to enhance, not restrict, citizens' ability to move and participate (Aditya et al., 2023; Ji et al., 2021). Thus, understanding how digital inequalities manifest in public transport is critical for designing more inclusive and equitable mobility systems.

Beyond digital inequality theorised in the digital divide framework, existing research has increasingly argued for a broader framework of digital justice (Cardullo et al., 2019; Heeks, 2021; Shelton et al., 2015), being the fair distribution of digital opportunities, representations, and decision-making authority within urban systems. Digital justice shifts the focus from whether citizens have access to digital tools toward how these tools reinforce or mitigate existing social inequities. In the context of mobility and public transport, the capacity to access, understand, and benefit from digital public transport services becomes a prerequisite for widened participation in the digital city. This perspective aligns closely with emerging work on urban equity and mobility justice, which emphasise that equitable mobility is assisted by not only physical infrastructure but also on digital accessibility (Ayre, 2020; Guzman & Bocarejo, 2017; Lucas, 2012; Verlinghieri & Schwanen, 2020).

Previous research on digital divide in public transport has primarily relied on surveys (Durand et al., 2024), interviews (Durand, Zijlstra, Hamersma, Hoen, et al., 2023), or focus groups(Graham et al., 2018; Owen et al., 2012), offering valuable qualitative insights but often limited in scope and temporal coverage. While such approaches reveal perceived barriers and motivations, they are typically constrained by small sample sizes, recall bias, and geographic focus. In contrast, the increasing availability of user-generated online data, such as mobile app reviews, provides an opportunity to capture naturally occurring feedback that reflects authentic and unprompted experiences (Chen et al., 2017; Mahmud et al., 2022; Verma, 2022). These data offer a valuable complement to traditional research, enabling continuous monitoring of user sentiment, behaviour, and expectations toward digital public services. By linking digital divide theory with justice-oriented frameworks, the current study conceptualises digital inequity in public transport as a form of urban injustice that can be empirically traced through digital expressions of frustration, adaptation, and demand for fairer services.

The application of Natural Language Processing (NLP) methods (Çapali et al., 2023; Chowdhary, 2020), including topic modelling (e.g., Roberts et al., 2014) and sentiment analysis (e.g., Gloutnikov, 2018; Verma, 2022), provides a data-driven, systematic, and efficient way to extract patterns from unstructured textual data. NLP is an artificial intelligence (AI) approach that enables the identification of latent themes, linguistic cues, and emotional tones that reflect users' lived experiences with digital technologies (DeNardis & Hackl, 2015). Topic modelling uncovers recurrent and overlapping topics to reflect hidden structures among texts, while sentiment analysis captures the polarity and intensity of opinions, whether positive, neutral, or negative, revealing how citizens emotionally respond to digital systems (Ma et al., 2025; Mahmud et al., 2022; Pineda-Jaramillo et al., 2023). Within the governance and planning literature, these methods have been recognised as powerful tools for "listening" to public discourse at scale, offering insights into public perceptions that can inform responsive and evidence-based policy design (Chen et al., 2017; Helsper, 2021). Despite this potential, few studies have applied NLP to examine digital inequality in public transport services.

Most existing NLP work has focused on consumer satisfaction or brand management in commercial contexts (Hirschberg & Manning, 2015), leaving a gap in understanding how digital divides shape interactions with public infrastructure. In transport research, sentiment analysis has primarily been used to evaluate travel experience or service reliability (Mahmud et al., 2022; Pineda-Jaramillo et al., 2023), while the social dimensions of digital participation, such as inclusion, accessibility, and trust, remain underexplored. Moreover, as smart mobility increasingly becomes a proxy for digital citizenship (Wilson et al., 2019), assessing who benefits and who is excluded from these systems becomes central to both urban planning and social equity agendas.

This study contributes to filling this gap by integrating NLP with digital divide theory to analyse how users engage with public transport applications in Birmingham as a case. Specifically, we examine user reviews from the Apple App Store and Google Play Store for six major transport services endorsed by the Birmingham City Council. By analysing over 4000 reviews collected between 2016 and 2025, we address three key research questions (RQ):

(1) What are the main topics and themes discussed in user feedback about public transport apps?
(2) How do users' emotional responses, captured through sentiment analysis, reflect satisfaction or frustration with these digital services?
(3) How do these themes and sentiments map onto the three levels of the digital divide (access, skills, and outcomes)?

By leveraging data-driven NLP text analysis of user-generated data, this study offers both theoretical and practical contributions. Theoretically, it extends digital divide research into the field of smart mobility, by linking the three levels of digital inequality to specific aspects of user interaction with public transport technologies. Practically, it demonstrates how citizen-generated data can inform more inclusive and equitable digital service design, providing researchers, policymakers, or stakeholders with actionable evidence about the intervention areas in digital mobility systems (Ojadi et al., 2023). Note that, the present study focuses on citizens who engage with official public transport applications and therefore examines inequalities in users' digital experiences rather than absolute digital exclusion. Ultimately, understanding how citizens perceive and emotionally evaluate digital public transport services is not only a question of technological efficiency but also of urban equity and digital justice. In cities such as Birmingham, where disparities in both digital competence and transport access persist, such insights are crucial for ensuring that digital transformation advances, rather than undermines, the principles of accessibility and fairness at the core of sustainable urban development.

2. Methodology

2.1. Data collection

We focused on public transport services and their officially recognised mobile applications (i.e., Apps) in Birmingham, as recommended by the Birmingham City Council. These services included six public transport mobile applications: West Midlands Railway (trains), West Midlands Metro (trams), NX Bus (buses), TfWM/Swift (trains, trams, buses), Swift Collector (trains, trams, buses), Beryl (city bikes and e-scooters). We chose these six mobile applications as they were officially developed and endorsed by local authorities for public service. Third-party applications (e.g., Trainline) were excluded because the aim of this study was to understand citizen participation in government-provided digital services rather than private platforms. Anonymous user reviews of these six public transport mobile applications (Fig. 1) were extracted from both the Apple App Store and Google Play Store. User reviews were retrieved between September 2016 and October 2025 using a custom Python web-scraping script (data retrieved on 18 October 2025) with the *beautifulsoup* library (Richardson, 2007). The 10-year period corresponded to the full operational lifespan of these applications, from the earliest available review to the most recent at the time of data collection for each application. The dataset included user reviews across multiple mobility services (e.g., trains, buses, trams, city bikes, and e-scooters) and spanned a 10-year period corresponding to the operational lifespan of these applications. While these services differed in specific user contexts, all applications represented an integrated interface of official digital platforms, through which citizens interact with public transport services. Our data and subsequent analyses therefore focused on common patterns in citizen feedback regarding public transport digital services as a whole. Aggregating reviews across applications and platforms also allowed the construction of a larger and more diverse text corpus, which could improve the robustness of NLP-based analyses by enabling recurring themes and sentiment patterns to emerge more reliably.

The dataset included the publicly available information: app name, platform, review date, users' star rating (1 = worse, 5 = best), and the free-text review. To ensure comparability, and as the focus of the current study was the city of Birmingham, only English-language reviews from the UK versions of the app stores were retained, for both Apple App Store and Google Play Store. To ensure the completeness of the dataset, only data entries with both ratings and free-text reviews were included. Because the two platforms differ in data structure (e.g., the App Store includes review titles, whereas Google Play does not), we restricted analysis to fields that were structurally consistent across both platforms to maintain consistency and balance across the dataset.

Our data collection approach was taken to maintain the integrity of the dataset and ensure that even less popular public transport services were adequately represented. By doing so, we aimed to capture a comprehensive range of user experiences and sentiments across various public transport applications in Birmingham. We strictly adhered to ethical guidelines to ensure the privacy and anonymity of individual users: no personal identifiable information (user IDs, names, IP address, locations) were collected. The data used for analysis consisted solely of publicly available information. All data were publicly available, anonymised, and used solely to examine aggregate user experience and sentiment toward public transport applications. This study was approved by the local Ethics Committee at the Birmingham City University (Ethics approval ID: 12628).

2.2. Cleaning and preprocessing

All text was cleaned and standardized prior to analysis in R with

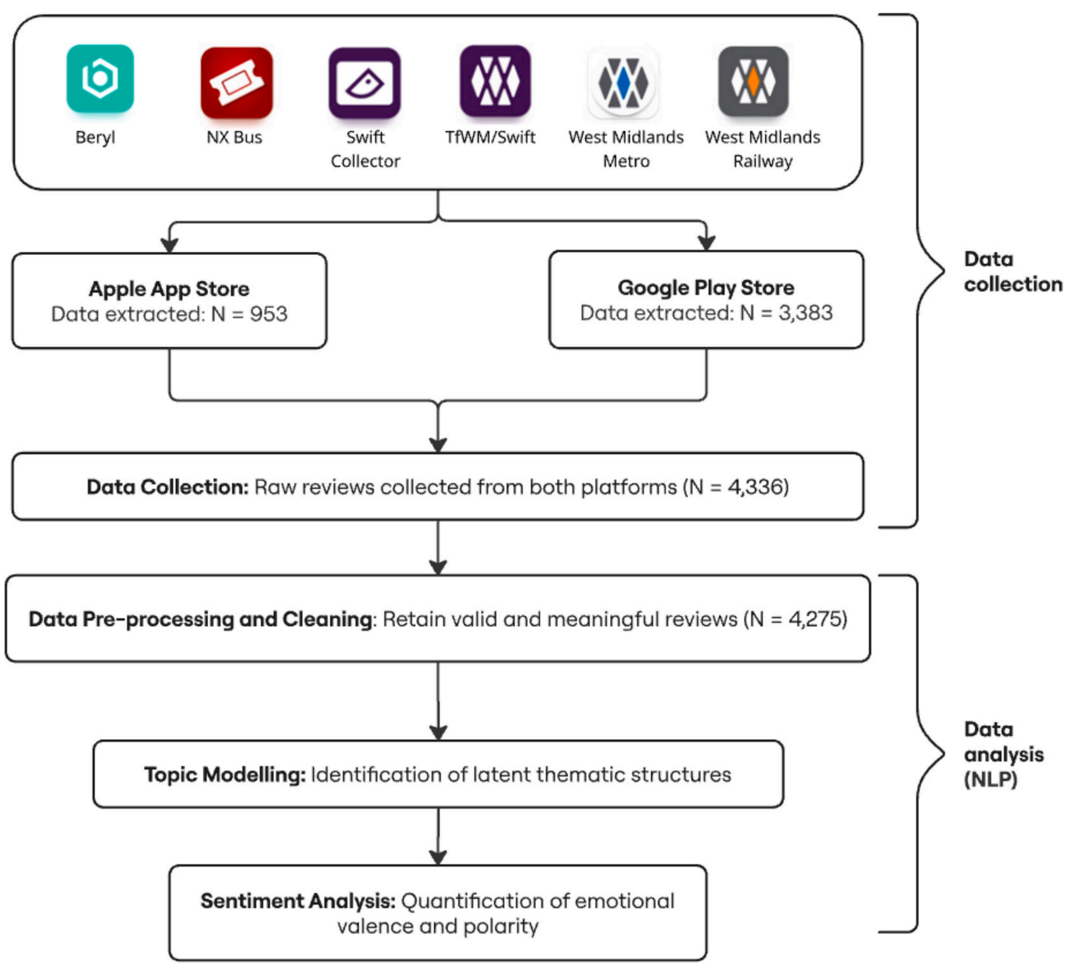

Fig. 1. Analytical pipeline. Raw User reviews from six mobile public transport applications were extracted from Apple App Store and Google Play Store. After data cleaning and processing, topic modelling was used to identified latent themes, followed by sentiment analysis quantifying emotional valence and polarity toward these applications.

appropriate packages (Table 1). These included: (1) Duplication removal: removing duplicate entries based on identical app name, date, rating, and review text; (2) Normalization: lower-casing all texts, removing punctuation, numbers, URLs, and special symbols (e.g., @, #), and converting emojis to plain texts with the *texclean* and *stringi* package; (3) Lemmatization: words were Lemmatized to their dictionary form using the *textstem* package, to account for linguistic variants (e.g., purchased ➔ purchase). To preserve semantically important adjectives, an exception list of words (e.g., best, worst, faster, amazing) were protected from lemmatization. (4) Stop-word removal: removing standard English stop words (e.g., "and", "the", "this", etc.) and a small set of domain-specific high-frequency words (e.g., "app", "apps") and application names (e.g., "beryl") that carried little semantic value under the context of investigation; (5) Tokenization: splitting reviews into word tokens, namely, smaller, manageable, and meaningful units like words for NLP analyses, using the *quanteda* package. Tokens shorter than three characters were removed. Reviews that became empty after cleaning were flagged and excluded from model fitting. Each review was assigned a document identifier ("doc_id"), which allowed all topic, sentiment, and metadata to be re-merged consistently.

2.3. Topic modelling

To uncover latent patterns such as overlapping topics from unstructured data as in mobile app reviews, we performed topic modelling to tokenised data resulted from data preprocessing. Because user reviews of mobile apps are short and concise (e.g., often one to two sentences) as opposed to journal articles, blogs, or books that are long and comprehensive, standard topic modelling approaches such as Latent Dirichlet Allocation (LDA; Blei et al., 2003), Correlated Topic Models (CTM; Blei & Lafferty, 2007), or Structural Topic Models (Roberts et al., 2014) are less suitable. We therefore employed Biterm Topic Modelling (BTM; Yan et al., 2013), a reliable topic modelling approach most suitable for short and sparse texts. Unlike conventional topic models such as LDA, which estimate topic mixtures separately for each document, BTM directly models the co-occurrence of word pairs (biterms) across the entire corpus. Although the text was tokenised into individual words during preprocessing, BTM does not analyse words independently. Instead, it extracts and models frequently co-occurring word pairs, allowing common expressions and semantic relationships to jointly inform topic estimation. By borrowing information from global word co-occurrence patterns rather than relying on individual documents, BTM produces more stable and coherent topic estimates when documents are short and contain relatively few words, making it particularly well suited to mobile application reviews. Consequently, the inferred topics reflect patterns of word co-occurrence rather than isolated lexical items. BTM was fitted with the *BTM* package in R. Each review's tokens were converted into a two-column data frame (doc_id, token), representing all word occurrences per review. For the current study, we used the BTM parameters following the best BTM practice (Yan et al., 2013): number of topics $K = 3$, Dirichlet prior probability for topic distribution $\alpha = 50/K$, Dirichlet prior probability for word distribution given the topic $\beta = 0.01$, and number of iterations of Gibbs sampling $I = 2000$. To optimise the value of K, a grid search was performed across $K = \{3, 4, 5, 6\}$, and we assessed BTM model performance for each K value using topic coherence, topic distinctness, and topic concentration diagnostics (Yan et al., 2013). A composite BTP diagnostic score was computed using the following equation: Composite score $= 0.45 * \text{coherence} + 0.45 * \text{distinctness} + 0.10 * (1 - \text{concentration})$. Higher composite score indicated more meaningful K. For each topic, BTM estimated a topic-word probability distribution, from which, the top 50 most probable words were extracted for visualization. Note that BTM allowed mixed membership (i.e., a word can contribute to multiple topics), providing greater flexibility in identifying latent meanings within short, unstructured text.

Table 1
Summary of data analysis pipeline.

	Analysis pipeline	Software Package
(A) Data preparation	(A1) Duplication removal	Base R
	(A2) Basic normalization	*texclean, stringi*
	(A3) Lemmatization	*textstem*
	(A4) Stop-word removal	Base R, *quanteda*
(B) Topic modelling	(B1) Tokenisation	*quanteda*
	(B2) Biterm topic modelling	*BTM*
(C) Sentiment analysis	(C1) Dictionary-based sentiment analysis	*sentimentr*
	(C2) Rule-based sentiment analysis	*VADER*

2.4. Sentiment analysis

To capture the sentiment, namely the emotional tone, or the valence of subjective opinion within the users' reviews, we additionally conducted sentiment analysis after topic modelling, with a hybrid approach combining the dictionary-based sentiment analysis implemented in the *sentimentr* package in R and rule-based sentiment analysis implemented in the *VADER* (Valence Aware Dictionary and sEntiment Reasoner) package in R. The *sentimentr* package (Rinker, 2016) computed sentence-level valence (e.g., positive vs. negative) while accounting for local negations (e.g., not helpful) and intensifiers (e.g., very bad), whereas *VADER* (Hutto & Gilbert, 2014) incorporated punctuation, capitalization, and emoji cues that are specifically attuned to short and informal texts such as application reviews. Both approaches were applied to the lemmatised text using the same preprocessing pipeline as in topic modelling except that stop words were removed to reflect instances such as negation. In addition, prior to sentiment analysis, all reviews were independently screened by two raters for the presence of irony or sarcasm. Reviews judged by both raters to contain ironic or sarcastic expressions were adjusted, with any disagreements resolved through discussion and consensus. Following this screening, a small number of domain-specific lexical adjustments (e.g., assigning negative polarity to terms such as "joke" and "laughable") were introduced to improve the interpretation of recurrent transport-related expressions. These adjustments were intended to improve domain sensitivity rather than to detect irony itself. Finally, sentiment scores obtained from both approaches were z-standardized and averaged to produce a composite "ensemble" sentiment score to reliably reflect the overall sentiment per user reviews. Using the ensemble sentiment score, user reviews were then categorised as positive (greater or equal to 0.05), negative (smaller or equal to -0.05), or neutral (between -0.05 and 0.05). This is to provide a buffer zone around zero to minimise classification noise. The ensemble sentiment scores were additionally correlated with the original star rating to provide further validation of the sentiment analysis approach. Finally, we sought to bridge results from both topic modelling and sentiment analysis at the review level, namely, to examine which words contributed most to the sentiments under each topic per user review. Topic-level sentiment was computed as the average ensemble sentiment of reviews that were confidently assigned to each topic (i.e., posterior probability ≥ 0.30 after BTM), and log-likelihood ratios (χ^2) were computed to identify words most associated with positive versus negative sentiment within each topic.

3. Results

3.1. Descriptive statistics

A total of 4336 user reviews were extracted from web scraping. After data cleaning and preprocessing, the final dataset contained 4275 user reviews, comprising 953 reviews from the Apple App Store and 3322 from the Google Play Store (Table 2). The average review length after cleaning and preprocessing was 24.81 words (Standard deviation, SD =

Table 2
Overview of data.

	Apple App Store			Google Play Store		
	# of records	Start	End	# of records	Start	End
Beryl	355	Jun. 2019	Oct. 2025	838	Mar. 2019	Oct. 2025
NX Bus	287	Oct. 2016	Sep. 2025	1237	Sep. 2016	Sep. 2025
Swift Collector	57	Mar. 2020	Sep. 2025	251	Aug. 2017	Oct. 2025
TfWM/Swift	79	Nov. 2021	Aug. 2025	316	Mar. 2021	Sep. 2025
West Midlands Metro	34	Jan. 2019	May. 2025	218	Jan. 2019	Oct. 2025
West Midlands Railway	141	Dec. 2017	Sep. 2025	462	Dec. 2017	Oct. 2025
Total	953			3322		

23.66). Review lengths were positively skewed (median = 18 words, Interquartile Range, IQR between 1st and 3rd quantile = 8–34 words). 9.9% of reviews contained three words or fewer. Reviews without accompanying free-text comments were excluded before analysis. Although many reviews were relatively brief, this is characteristic of mobile application data, and the Biterm Topic Model employed in the present study is specifically designed for short, sparse texts (see Section 3.2 below). The average user star rating was 2.06 (SD = 1.54), which was statistically significantly lower than three on a five-star scale (the theoretical neutral score between one and five would be three stars; one-sample *t*-test, $t_{4274} = -39.79$, $p < 0.001$), indicating generally negative user evaluations across applications. Fig. 2 shows the temporal distribution of reviews and ratings per application between 2016 and 2025. This period covered the time between the app's induction and data collection of the current study, hence was representative of their “lifetime” development.

3.2. Topic modelling

To answer our RQ1 on “What are the main topics and themes discussed in user feedback about public transport apps?” and RQ3 on “How do these themes and sentiments map onto the three levels of the digital divide (access, skills, and outcomes)?”, Biterm Topic Modelling (BTM) was used to identify recurrent topics and latent structural patterns across user reviews. Three topics emerged from the user reviews, as model diagnostics indicated that $K = 3$ offered the best balance among topic coherence, topic distinctness, and topic concentration (Table 3; Fig. 3B), suggesting interpretable and non-redundant topics. For completeness, we also examined higher-topic solutions ($K = 4$–6), which primarily subdivided the identified themes into more specific subtopics without revealing additional conceptually distinct dimensions, and were therefore considered less parsimonious for the present objectives.

Each topic was represented by its most probable words and summarised by an interpretive label assigned based on the semantic content (Fig. 3A). Topic 1 was labelled as “App performance” (relative prevalence of 48.3% among all user reviews; Fig. 3C), and top words under this topic were commonly associated with registering, logging-into, and payment. Topic 2 was labelled as “Usability” (relative prevalence of 22.3%; Fig. 3C), and top words under this topic were commonly associated with registering with usefulness, functionality, and readability of information. Topic 3 was labelled as “Satisfaction” (relative prevalence of 29.3%; Fig. 3C), and top words under this topic were commonly associated with registering with customer service, cost, and refund.

Importantly, the identified topics can be interpreted and mapped onto the three-level digital divide framework, namely access, ability, and motivation. Topic 1 keywords related to registration and payment primarily reflected access-level barriers, indicating technical or infrastructural constraints to accessing digital services. Topic 2 keywords

Table 3
Topic number optimisation.

K	coherence	distinctness	concentration	Composite score*
3	−2.18	0.78	0.31	0.54
4	−2.16	0.78	0.34	0.51
5	−2.32	0.79	0.31	0.31
6	−2.39	0.84	0.38	0.45

* Composite scores = 0.45 * coherence +0.45 * distinctness +0.10 * (1 – concentration) to balance the three diagnostic characteristics.

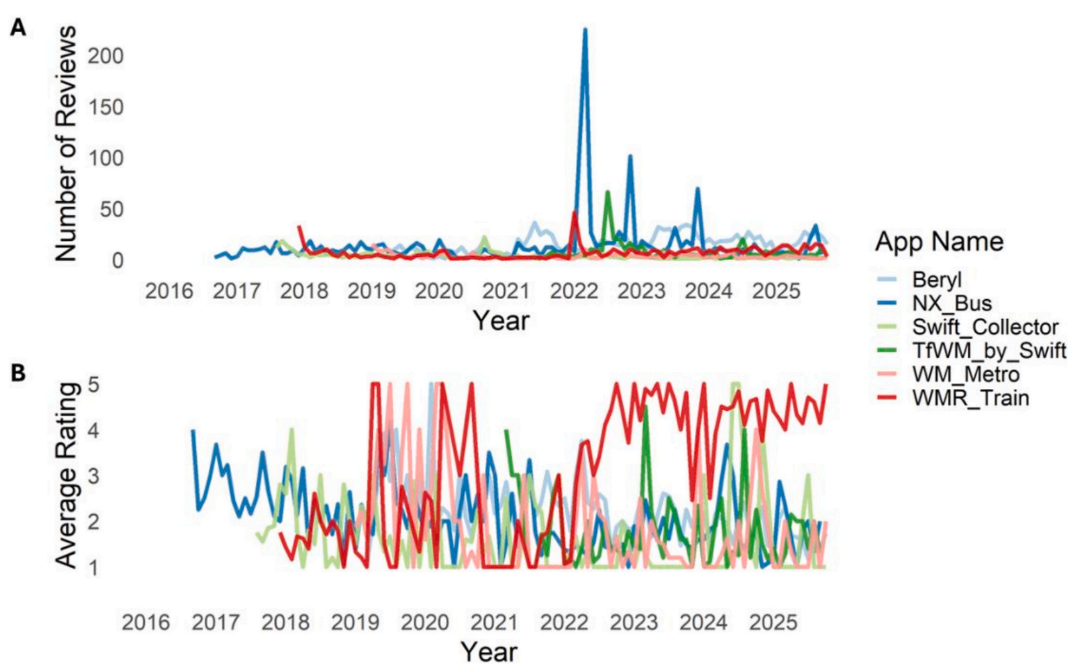

Fig. 2. Temporal distributions of (A) number of reviews and (B) average star ratings across time.

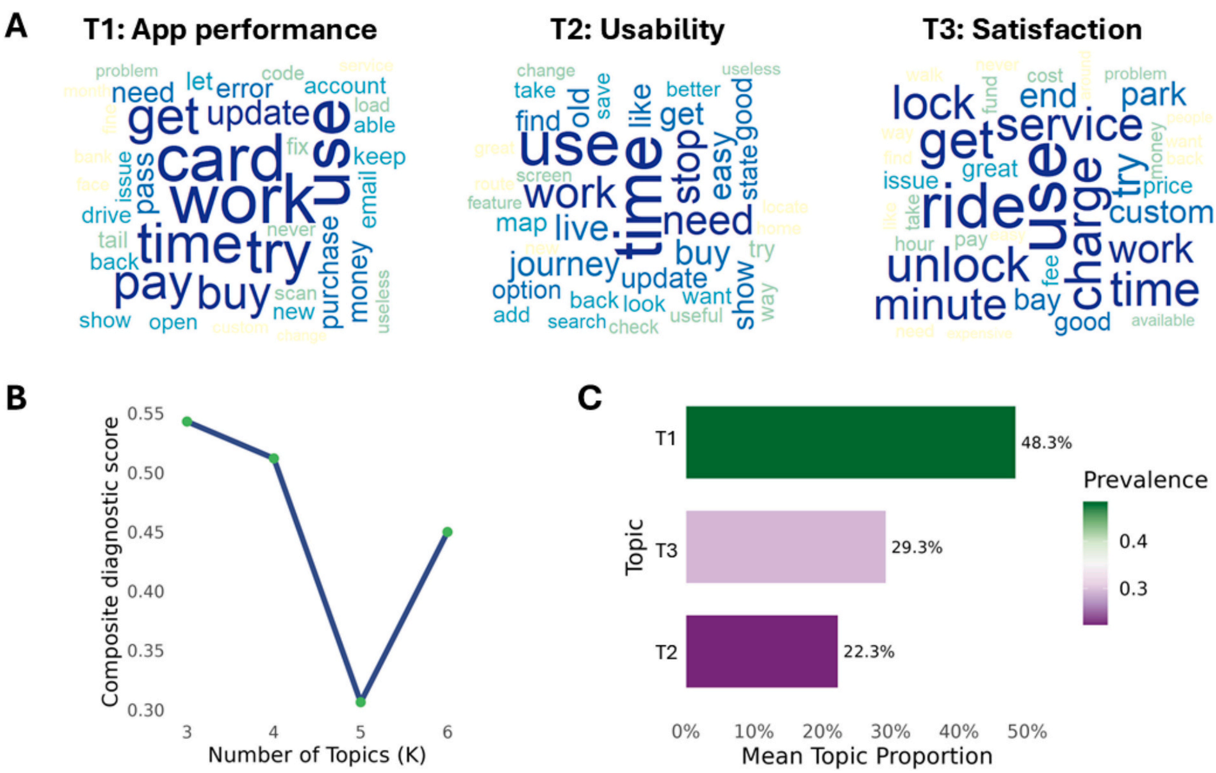

Fig. 3. Topic modelling. (A) Word cloud showing most probable words for each topic. (B) Topic number identification balancing among topic coherence, topic distinctness, and topic concentration. A higher composite diagnostic score indicates the suggested number of topics (K). Scatters indicate user rating and ensemble sentiment scores, and the redline shows the regression line between them. (C) Topic prevalence shown as overall proportions.

related to usefulness and functionality corresponded to ability-level differences, revealing how users' digital literacy, confidence, or support availability affected their experience. Topic 3 keywords related to service, support availability, and refund, in contrast, mapped onto the motivational level, where positive experiences enhanced continued use and trust in digital public transport services. Together, this theory-grounded integrative mapping suggested that user experiences captured in app reviews provide nuanced insight into how digital inequality manifests in everyday interactions with public transport services in the digital world.

3.3. Sentiment analysis

To answer our RQ2 on "How do users' emotional responses, captured through sentiment analysis, reflect satisfaction or frustration with these digital services?", sentiment analysis was used to capture users' tones and valence when expressing opinions on the public transport apps. Sentiment scores were computed using the hybrid ensemble approach combining *sentimentr* and *VADER* on the user reviews, creating an ensemble sentiment score (see Methods). The two sentiment analysis packages showed strong positive correlation ($r = 0.598$, $p < 0.0001$), confirming consistent valence detection across methods. Across the user reviews, 1950 of them (45.61%) were classified as positive, 2151 as negative (50.32%), and 174 (4.07%) as neutral (Fig. 4A), showing that most reviews were polarised with only a handful of neutral comments. The ensemble sentiment score was positively correlated with users' original star ratings ($r = 0.553$, $p < 0.0001$; Fig. 4B), meaning that higher ensemble sentiment scores were associated with higher users' original star ratings. This positive correlation further validated the measure's reliability of our sentiment analysis approach.

In addition, ensemble sentiment scores varied across both applications and platforms (Table 4), indicating heterogeneity in user evaluations. Across apps, TfWM/Swift showed the most negative sentiment (Apple: −0.57; Google Play: −0.25), followed by West Midlands Metro and NX Bus, which also exhibited negative mean scores on both platforms. Swift Collector displayed platform-specific divergence, with near-neutral sentiment on the Apple App Store (0.02) but more negative evaluations on Google Play (−0.14). In contrast, West Midlands Railway showed consistently positive sentiment across both platforms (Apple: 0.19; Google Play: 0.51), while Beryl remained close to neutral overall, with a slight negative mean on Apple (−0.10) and positive mean on Google Play (0.06). These results demonstrated systematic variation across apps and platforms, alongside the overall distribution of sentiment being slightly negative when aggregated across the full dataset.

3.4. Topic-sentiment interaction

The sentiment analysis revealed the overall tone across all user reviews, and now that we identified three topics from the topic modelling (Section 3.2), to answer our RQ3 on "How do these themes and sentiments map onto the three levels of the digital divide (access, skills, and outcomes)?",

we sought to connect topics and sentiments by examining the topic-sentiment interaction. Topic-level sentiment was estimated as the mean ensemble sentiment of reviews confidently assigned to each topic. Namely, we only focused on those review comments that received at least a 30% probability of being associated with a topic from topic modelling, and then examined the corresponding ensemble sentiment score for the associated topic. For example, if a user review entry was identified to have a 68% probability of belonging to Topic 1, 23% of Topic 2, and 9% of Topic 3, we assigned the label of Topic 1 to this review entry, and then connected the corresponding ensemble sentiment score with Topic 1. On average, the topic-level sentiments were − 0.17 for Topic 1, 0.29 for Topic 2, and 0.16 for Topic 3 (Fig. 4C), suggesting that users were generally positive about usability, but less so about satisfaction, and mostly negative about app performance. This gradient

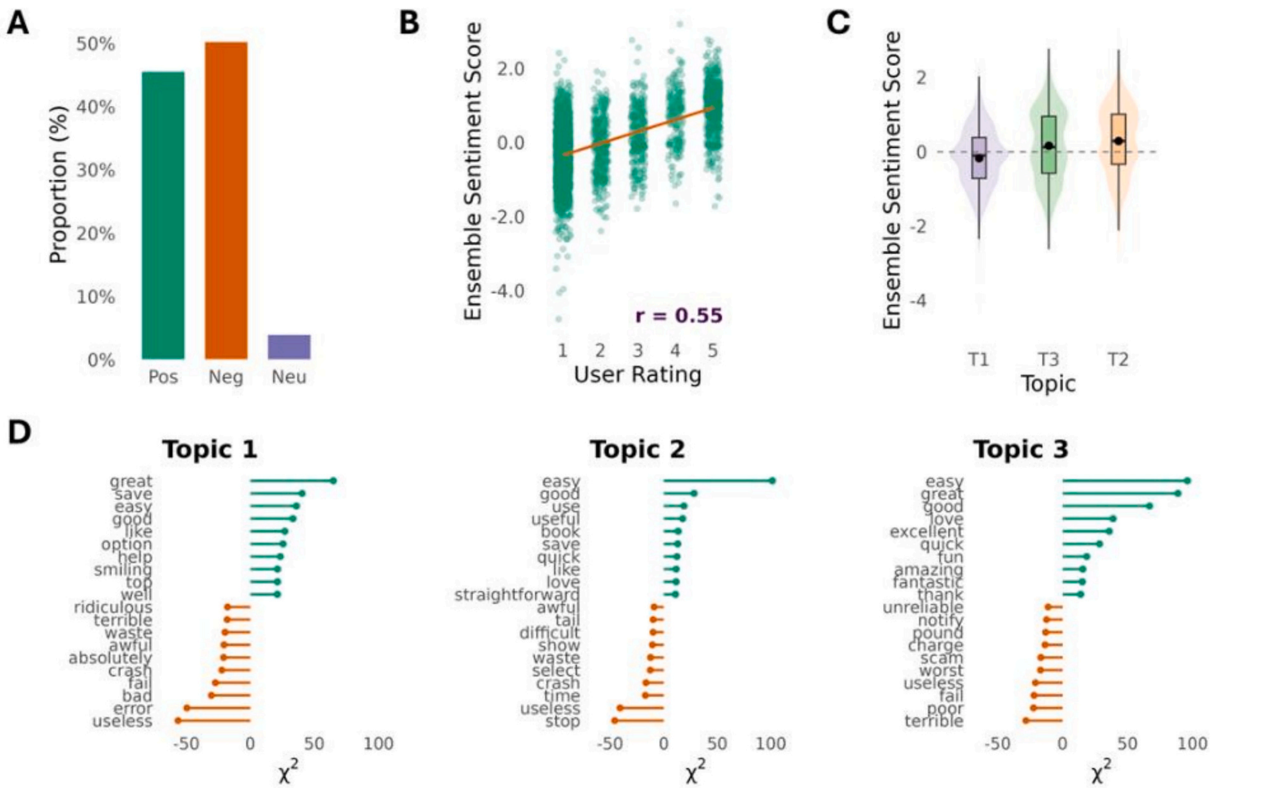

Fig. 4. Sentiment analysis. (A) Proportion of sentiments across user reviews. Pos = Positive, Neg = Negative, Neu = Neutral. (B) Association between ensemble sentiment score and original user star ratings. Ensemble sentiment scores are computed by combining the *sentimentr* and *VADER* approaches to reliably reflect the overall sentiment. With the ensemble sentiment score, user reviews are categorised as positive (greater or equal to 0.05), negative (smaller or equal to −0.05), or neutral (between −0.05 and 0.05). (C) Average topic-specific sentiments shown as violin plots. The black dot shows the mean, the edges of the box show the upper and lower 25th and 75th percentile, respectively, and the violin shadings indicate the data density. (D) Lexical markers associated with positive versus negative reviews. These lexical markers provide an interpretable summary of review-level statistically associated sentiment but should not be interpreted as independent semantic units, as many occur within larger multi-word expressions.

Table 4
Ensemble sentiment scores across application and platform.

	Ensemble sentiment: mean ± sd*		
	Average	Apple App Store	Google Play Store
Beryl	0.01 ± 0.93	−0.10 ± 0.96	0.06 ± 0.92
NX Bus	−0.07 ± 0.81	−0.03 ± 0.82	−0.08 ± 0.81
Swift Collector	−0.11 ± 0.90	0.02 ± 0.97	−0.14 ± 0.88
TfWM/Swift	−0.31 ± 0.88	−0.57 ± 0.84	−0.25 ± 0.87
West Midlands Metro	−0.05 ± 0.84	−0.16 ± 0.80	−0.04 ± 0.85
West Midlands Railway	0.43 ± 0.90	0.19 ± 1.00	0.51 ± 0.85

* sd = standard deviation.

showed distinct and asymmetric distribution of users' sentiments on the three identified topics. To further characterise sentiment polarity, namely which words were contributing to the positive vs. negative valence of the topic-level sentiment, log-likelihood ratios (χ^2) were computed for words most associated with positive versus negative sentiment within each topic. Lollipop plots (Fig. 4D) highlighted distinct lexical markers of satisfaction (useful, helpful, or quick, easy) versus frustration (crash, error, difficult, or unreliable) across topics. These associations reflected review-level sentiment according to statistical association rather than the inherent polarity of individual words.

4. Discussion

The current study examined how the digital divide manifested within Birmingham's public transport ecosystem by analysing more than 4000 user-generated textual reviews collected from the Apple App Store and Google Play Store. Using NLP with a combination of topic modelling and sentiment analysis, we assessed how users perceive, describe, and emotionally evaluate their experiences with digital mobility services. Our analyses revealed three dominant topics, app Performance, usability, and service satisfaction, which map closely onto the three levels of the digital divide framework (Durand et al., 2022; Helsper, 2021; Van Deursen & Helsper, 2015). Sentiment analysis further showed that users' overall evaluation of public transport apps was negative on average, particularly in relation to app performance and customer service. Together, these findings extended the digital divide literature to the field of smart mobility and demonstrated how AI-informed NLP can provide empirically grounded evidence for evaluating the inclusiveness of digital public services, showcasing digital twin approaches for sustainable and efficient public transportation (Manandhar et al., 2025).

4.1. Promise of NLP for digital inequality research in urban planning

Within urban and transport planning, the study of digital inequality has traditionally relied on surveys, interviews, or focus groups to capture people's attitudes toward technology (Durand et al., 2024; Durand, Zijlstra, Hamersma, Hoen, et al., 2023; Graham et al., 2018; Owen et al., 2012). While such approaches remain valuable for in-depth understanding, they are often limited in scale, temporal coverage, and representativeness. By contrast, publicly available data such as user reviews entail a naturally occurring source that captures spontaneous and authentic feedback on a focused topic from a diverse range of users over time (Chen et al., 2017; Mahmud et al., 2022; Verma, 2022). In this study, the rich data of more than 4000 reviews, spanning 2016 to 2025, allowed us to examine not only the evolution of user engagement and star ratings but also the recurring topics and the underlying semantic and emotional patterns embedded in those topics, which offered insights into common themes around digital mobility and the corresponding emotional attitude toward them.

By applying NLP methods, we were able to uncover latent structures in the textural data that would not have been detectable through conventional approaches, such as survey-based text analysis, thematic analysis, or keyword frequency counts (Aldabbas et al., 2020; Çapali et al., 2023; Feuerriegel et al., 2025; Hirschberg & Manning, 2015; Ma et al., 2025). Topic modelling identified the main dimensions of user reviews, while sentiment analysis provided quantitative evidence of the affective tone associated with each dimension. Importantly, we followed best open science practices: all analyses were conducted using open-source software in R and Python and corresponding packages, and the complete code pipeline has been made publicly available to promote transparency and replicability (Collaboration, O. S, 2015; Jin et al., 2023; Vicente-Saez & Martinez-Fuentes, 2018). This open-science approach strengthens the methodological credibility and rigour in the research of smart city planning and the analytical pipeline can serve as a template for future studies seeking to integrate user-generated data into urban governance (e.g., Yap et al., 2022). Similar data-driven approaches have been adopted in digital infrastructures (Cavallaro & Dianin, 2022), disaster management (Ma et al., 2025) and public health (Liu et al., 2019), but such application to digital inclusion in mobility systems remains novel. Note that, for our sentiment analysis, although manual screen was employed to reduce the influence of irony and sarcasm on sentiment estimation, this procedure inevitably involves human judgement and may require adaptation for different application domains. Future studies could evaluate automated context-aware approaches, including large language models, for scalable irony detection while maintaining transparent validation against human annotations.

4.2. Common topics linking three levels of the digital divide theory in smart mobility

Our identified three main topics, *App Performance*, *Usability*, and *Service Satisfaction,* offered a meaningful conceptual alignment with the three levels of the digital divide, and extended the digital divide theoretical framework into the field of smart mobility, by examining specific aspects of user interaction with digital public transport services. Importantly, the identified topics were derived using an unsupervised topic modelling approach, and no a priori digital divide-related categories or keywords were specified or seeded in the analysis, hence offering unbiased data-driven insight. The mapping of topics onto the three-level digital divide framework should therefore be understood as an interpretive, theory-guided alignment rather than a deterministic or confirmatory classification that the digital divide necessarily comprises exactly three dimensions. Alternative topic solutions provided finer-grained subdivisions of these themes but did not fundamentally alter the overarching conceptual structure.

The *App Performance* topic, characterised by keywords such as “registration”, “login”, and “payment”, was interpreted as reflecting functional access barriers, as these issues concerned users' ability to successfully access and interact with the digital system in the first place, reflecting user complaints about issues commonly associated with login failures, account errors, and payment issues, etc. These corresponded to the first-level divide, or access divide, representing challenges in the technical or infrastructural access to use digital services. Prior studies in transport equity have identified similar access constraints, especially in low-income, elderly, or digitally excluded communities, where unreliable digital infrastructure limits participation in smart mobility initiatives (Van Deursen & Helsper, 2015; Helsper, 2021; Chang et al., 2021; Liu et al., 2025).

The *Usability* topic, associated with keywords such as “usefulness”, “navigation”, and “functionality”, was interpreted as reflecting ability-level differences, as these related to users' capacity to effectively navigate, use, and understand the application interface, capturing challenges such as software crash, inconsistent navigation, or unstraightforward functionality. This aligned with the second-level divide, or skills divide, where users' digital literacy and confidence shape their capacity to engage effectively with public transport applications. Even with equal access to smartphones and the internet, differences in digital competence can produce inequality (Durand et al., 2024; Durand, Zijlstra, Hamersma, Hoen, et al., 2023). Complex functionality and poor interface design could worsen these differences, making the digital service unintentionally biased toward more digitally experienced users.

The *Service Satisfaction* topic, characterised by keywords such as “customer service”, “service reliability”, and “refund”, was interpreted as reflecting outcome-level differences, encompassing issues related to customer service, refunds, and ticket validity. This topic hence aligned with the third-level divide or outcome divide (also termed as motivation divide), as these captured the benefits users derive from digital engagement and their willingness to continue using the service. This level concerned inequalities in the benefits people derive from digital participation (Ruiu et al., 2025; Van Deursen & Helsper, 2015). Even when users were able to access and operate the apps successfully, many expressed frustrations at not receiving fair or timely outcomes, such as delayed refunds or lack of problem resolution, leading to less motivation to further engage with these apps. This topic identification supported the notion that the digital divide may extend beyond access and skills to encompass differential returns from digital engagement.

Note that the present analyses were based on app-store reviews, the study necessarily reflects the experiences of individuals who had already downloaded and interacted with official transport applications. Consequently, the findings should not be interpreted as capturing the experiences of those who remain completely excluded from digital public transport services. Rather, they reveal how inequalities persist even after initial digital access has been achieved, through differences in usability, reliability, and the benefits users derive from digital engagement. These findings therefore complement, rather than replace, research on first-level digital exclusion. We acknowledge that these dimensions are not mutually exclusive and that some user experiences may span multiple levels of the digital divide. The present mapping should therefore be understood as a conceptual framework for interpreting emergent themes rather than a strict categorical separation.

4.3. Sentiment patterns and thematic insights of digital divide in public transport

Our sentiment analysis revealed an overall negative emotional tone across the user review, with more than half of all reviews classified as negative. Among the identified topics, *App Performance* displayed the most negative sentiment, followed by *Service Satisfaction*, while *Usability* was relatively more positive. This gradient (Fig. 3C) may suggest that app-related technical reliability and performance and service responsiveness and satisfaction, rather than user competence, may have been the most common sources of frustration and dissatisfaction, leading to more divided use of digital public transport services. The topic-sentiment analysis (e.g., lollipop plots) further highlighted key discriminative words contributing to positive and negative sentiment per topic. Words such as *easy*, *helpful*, and *efficient* were strong indicators of positive experiences, whereas *error*, *crash*, *refund*, and *useless* dominated negative feedback. These findings resonated with studies of user trust in digital mobility platforms (Mahmud et al., 2022; Pineda-Jaramillo et al., 2023), which similarly emphasise the role of reliability and service-oriented care in shaping digital satisfaction. From a planning perspective, such insights are valuable, because they revealed where the digital transformation of public services may inadvertently undermine transport equity, by creating new forms of exclusion tied to software design and performance rather than physical access.

While the present analysis focused on aggregated patterns across all applications, the dataset included app-level metadata that allowed for more fine-grained comparisons (Table 4). Such analyses examined whether sentiment systematically differed across individual applications. The findings indicated that, beyond variation at the level of functional themes, additional differences exist across apps and platforms. Several applications, including TfWM/Swift, NX Bus, and West Midlands Metro, exhibited consistently negative sentiment, likely in relation to app performance and service reliability, whereas West Midlands Railway showed positive evaluations across both platforms. Other services, such as Beryl and Swift Collector, displayed more neutral or platform-dependent sentiment patterns. For example, higher-usage apps such as NX Bus tended to attract more negative feedback, likely related to app performance and technical reliability, whereas other services showed relatively more balanced sentiment profiles. These differences may reflect variation in service complexity, user expectations, and usage frequency across digital public transport applications. These results suggest that while key challenges were broadly shared, their intensity varies across applications, highlighting the need for both system-wide improvements and targeted, app-specific interventions.

4.4. Linking digital divide, digital justice, and urban equity

Our findings illustrate that the digital divide is not only a purely technical phenomenon but a dimension of digital justice within the urban governance of mobility. Users' frustrations identified in topic-level sentiment analysis about malfunctioning apps, inaccessible interfaces, or unresponsive support systems reflect not only individual inconveniences but also systemic asymmetries in how digital services are designed, maintained, and serviced. When certain groups face greater barriers to accessing or benefiting from digital transport systems (Durand, Zijlstra, Hamersma, Hoen, et al., 2023), additional challenges around app performance, usability, and satisfaction compound existing mobility inequities, limiting opportunities for full participation in digital urban life. This aligns with broader calls for mobile equity and mobile justice (Cardullo et al., 2019; Guzman & Bocarejo, 2017; Heeks, 2021; Lucas, 2012; Shelton et al., 2015; Verlinghieri & Schwanen, 2020), which emphasise on transparency, inclusivity, and responsiveness in the deployment of smart services and technologies. Hence, addressing the digital divide in public transport calls for a combined action, to ensure that all citizens, not only those who are digitally literate or technologically competent, can access the benefits of smart mobility.

4.5. Limitations and future directions

Despite its innovation and contributions, several limitations merit consideration in the current study. First, the analysis focused solely on apps officially developed or endorsed by public transport providers, rather than commercial mobility apps (e.g., Trainline, UK Bus Checker). This focus aligned with the planning orientation of the study, evaluating top-down service provision, but it limited the breadth of comparison. Future research could contrast public and private platforms to explore how governance structures affect digital inclusivity.

Second, our analyses aggregated use reviews across applications and platforms which allowed the construction of a richer and more diverse text corpus regarding public transport digital services as a whole. Although the corpus contained over 4000 reviews spanning ten years, the number of reviews differed across applications and many individual reviews were relatively brief. While the BTM is well suited to short texts, richer textual sources or larger multi-city datasets may provide more detailed thematic resolution in future research. We also acknowledge that temporal changes in transport systems and app functionalities may shape user experiences and consider this context when interpreting the findings. For example, the substantially higher number of reviews for NX Bus on Google Play (Fig. 2A) appears to be associated with a software update that introduced faulty functionalities which triggered the sudden increase of user reviews. Therefore, future studies could further examine how user attitudes and behaviours toward digital public transport systems are temporally coupled with time-specific events. Analysing time-resolved semantic patterns around such events using advanced NLP approaches (e.g., dynamic topic models; Blei & Lafferty, 2006) together with recent developments in large language models (LLM; e.g., Sohail & Zhang, 2025) could provide more nuanced insights into how user concerns emerge and evolve over time. For example, LLMs provide promising alternatives by modelling contextual representations of entire sentences or documents rather than isolated lexical units. Such approaches may better capture nuanced expressions, including multi-word phrases, implicit meanings, and complex discourse structures. Nevertheless, LLM-based methods often involve reduced interpretability, greater computational demands, and challenges for reproducibility due to rapidly evolving model architectures and proprietary implementations. Future work may therefore benefit from combining transparent statistical NLP methods with context-sensitive LLM analyses to achieve both interpretability and richer semantic understanding.

Third, one might argue that the first-level divide, commonly defined as access to digital devices and the internet (Van Dijk, 2006), does not perfectly align with our *App Performance* topic, since app users by definition already possess these resources. We acknowledge this distinction but propose an expanded interpretation of “access” as functional digital access, encompassing the ability to reliably connect to, log into, and maintain interaction with digital systems. This redefinition captures the evolving nature of the digital divide in technologically mature societies. Indeed, it may further justify introducing a fourth-level divide, centred on interface design and user experience (UI/UX), aspects that have been commonly overlooked in previous studies (Martin-Sanromán et al., 2019). The UI/UX level concerns the user interface, offering a balanced combination between aesthetics and functionality, which is critical for strategizing comprehensive, accessible, and inclusive digital services even after initial access has been achieved. Future research could combine user-generated online data with surveys, interviews, or participatory research involving digitally excluded populations, older adults, or socioeconomically disadvantaged communities to provide a more comprehensive understanding of digital inequalities across all stages of digital engagement.

Forth, the study's geographical focus on Birmingham may seem to limit the results' generalisability. Indeed, Birmingham represents a theoretically rich case: it is the UK's second-largest city, socioeconomically diverse, and historically characterised by substantial digital exclusion (Allmann & Radu, 2023; Comber et al., 2022; Faith et al., 2022), which is more representative than other like-sized cities. Nonetheless, the use of anonymised online reviews precluded linking sentiment to demographic factors such as age, gender, or ethnicity, which are known dimensions relevant to digital divide. Future studies could triangulate our findings with in-depth participation research (Kopackova & Komarkova, 2020; Nurminen et al., 2024; Shtebunaev et al., 2023; Wilson et al., 2019) to link specific sentiment patterns to user characteristics, providing a more complete understanding of digital inequality in public transport services.

Last, while the present analysis identified overall patterns across mobile public transport applications and platforms, variation in user experiences may also reflect differences in user demographics associated with specific services. However, because the dataset consisted of anonymised app store reviews, it did not include demographic information such as age, gender, or socioeconomic background, as well as travel preferences. This limited the ability to directly examine how digital inequalities may vary across user groups. Future research could combine user-generated textual data with survey study or administrative data (Allard et al., 2018) to better understand how demographic factors shape digital engagement in public transport systems.

4.6. Implications for practice

Our findings using moderately sized, citizen-generated data may suggest implications for both urban planning and digital governance focusing on public transport services. Indeed, city planning strategies and policy suggestions have emphasised the importance of public transport in digital transformation globally, yet digital divide remains as a main challenge toward developing an inclusive and sustainable plan. (Durand, Zijlstra, Hamersma, 'T Hoen, et al., 2023; Garritsen et al., 2025; Nguyen & Hargittai, 2023). Beyond identifying patterns of user experience, the findings of this study highlight the potential of NLP-based analysis as a diagnostic tool to inform more responsive and user-centred digital public transport services. By systematically capturing feedback from 4275 users, this approach can help practitioners and policymakers identify areas of recurring dissatisfaction, particularly those related to technical performance and service responsiveness, and prioritise interventions accordingly.

Here the identified three themes and their sentiments highlight the distinct intervention areas to address challenges: It is important for policymakers and practitioners to be alerted that digital services are not "functional-by-default", and usability-related performance of these digital services and the satisfaction around using these digital services are equally critical. From the app performance perspective, public transport apps should be technically reliable and compatible across device platforms. From the usability perspective, public transport apps should include more straightforward design and reduce counterintuitive functionality. From the satisfaction perspective, public transport apps should require more responsive customer service and transparent feedback mechanisms. It is crucial for policymakers and practitioners alike to think carefully about both the integrated principles from which these services and systems are designed (Zukeran et al., 2024), and the mitigation strategies to alleviate unintended consequences in a more targeted fashion.

Importantly, these insights can serve as a foundation for more participatory and iterative service design. For example, integrating NLP-derived evidence with co-designed engagement approaches, such as focus groups involving customers, app developers, and transport providers, could enable a more direct translation of user feedback into service improvements. Such approaches, ideally combined with a longitudinal design, would support the development of more targeted and context-sensitive interventions, helping to align digital transport services more closely with user needs and expectations, and ultimately fostering more inclusive and effective digital mobility systems.

5. Conclusion

This study employed NLP of publicly available user-generated reviews to unveil public perceptions and sentiments of digital transport services and uncovered persistent forms of digital inequality in public transport systems. By linking computational text analysis with the three-level digital divide framework, we provided both methodological and theoretical contributions to the study of smart cities and transport equity. Our findings revealed that barriers to access, usability, and satisfaction hinder digital service delivery. Our open-source, reproducible pipeline exemplifies how data-driven tools such as NLP can be harnessed within planning and smart city research at scale. These insights underscore the need to design inclusive digital infrastructures that address not only connectivity and technical functionality, but also design interface, user confidence, feedback responsiveness, and perceived satisfaction. Integrating data-driven analytics with participatory approaches could further help ensure that the digital transformation of public transport enhances accessibility and trust.

CRediT author statement

L.H.: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing – Original Draft, Writing - Review & Editing, Visualization.

L.Z.: Methodology, Software, Validation, Formal analysis, Writing - Review & Editing.

S.G.: Writing - Review & Editing, Supervision.

M.H.T.: Writing - Review & Editing, Supervision, Funding acquisition.

CRediT authorship contribution statement

Lu Huafeng: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Data curation. **Lei Zhang:** Writing – review & editing, Validation, Software, Methodology, Formal analysis. **Silvia Gullino:** Writing – review & editing, Supervision. **Miguel Hincapié Triviño:** Writing – review & editing, Supervision, Funding acquisition.

Declaration of competing interest

All authors declare no conflict of interests.

Acknowledgement

This work was supported by a Dean's PhD scholarship 2022/2023 from the Birmingham City University to L.H. L.Z. was partially supported by the Wellcome Trust (228268/Z/23/Z), the Royal Society (IES\R3\243253), and the Birmingham–Nanjing University Joint Research and Education Centre pump-priming funding.

Data availability

Analysis code used in this study are publicly available on the Open Science Framework (OSF): https://osf.io/xbhuf/

References

Aditya, T., Ningrum, S., Nurasa, H., & Irawati, I. (2023). Community needs for the digital divide on the smart city policy. *Heliyon, 9*(8), Article e18932. https://doi.org/10.1016/j.heliyon.2023.e18932

Aldabbas, H., Bajahzar, A., Alruily, M., Qureshi, A. A., Amir Latif, R. M., & Farhan, M. (2020). Google play content scraping and knowledge engineering using natural language processing techniques with the analysis of user reviews. *Journal of Intelligent Systems, 30*(1), 192–208. https://doi.org/10.1515/jisys-2019-0197

Allard, S. W., Wiegand, E. R., Schlecht, C., Datta, A. R., Goerge, R. M., & Weigensberg, E. (2018). State agencies' use of administrative data for improved practice: Needs, challenges, and opportunities. *Public Administration Review, 78*(2), 240–250.

Allmann, K., & Radu, R. (2023). Digital footprints as barriers to accessing e-government services. *Global Policy, 14*(1), 84–94. https://doi.org/10.1111/1758-5899.13140

Ayre, L. B. (2020). What more can we do to address broadband inequity and digital poverty? *Information Technology and Libraries, 39*(3). https://doi.org/10.6017/ital.v39i3.12619

Blei, D. M., & Lafferty, J. D. (2006). Dynamic topic models. In *Proceedings of the 23rd International Conference on Machine Learning* (pp. 113–120).

Blei, D. M., & Lafferty, J. D. (2007). A correlated topic model of science. *The Annals of Applied Statistics, 1*(1). https://doi.org/10.1214/07-AOAS114

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research, 3*(Jan), 993–1022.

Çapali, B., KüçüksiLle, E., & Kemaloğlu Alagöz, N. (2023). A natural language processing framework for analyzing public transportation user satisfaction: A case study. *Journal of Innovative Transportation, 4*(1), 17–24. https://doi.org/10.53635/jit.1274928

Cardullo, P., Di Feliciantonio, C., & Kitchin, R. (2019). *The right to the smart city*. Emerald Publishing Limited.

Cavallaro, F., & Dianin, A. (2022). Combining transport and digital accessibilities in the evaluation of regional work opportunities. *Journal of Transport Geography, 98*, Article 103224. https://doi.org/10.1016/j.jtrangeo.2021.103224

Chang, J. E., Lai, A. Y., Gupta, A., Nguyen, A. M., Berry, C. A., & Shelley, D. R. (2021). Rapid transition to telehealth and the digital divide: Implications for primary care access and equity in a post-COVID era. *The Milbank Quarterly, 99*(2), 340–368. https://doi.org/10.1111/1468-0009.12509

Chen, N. C., Zhang, Y., Stephens, M., Nagakura, T., & Larson, K. (2017). Urban data mining with natural language processing: Social media as complementary tool for urban decision making. In *The 17th International Conference on computer aided architectural design futures* (pp. 101–109).

Chowdhary, K. (2020). Natural language processing. In *Fundamentals of Artificial Intelligence* (pp. 603–649).

Collaboration, O. S. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), Article aac4716.

Comber, S., Park, S., & Arribas-Bel, D. (2022). Dynamic-IMD (D-IMD): Introducing activity spaces to deprivation measurement in London, Birmingham and Liverpool. *Cities, 127*, Article 103733. https://doi.org/10.1016/j.cities.2022.103733

DeNardis, L., & Hackl, A. M. (2015). Internet governance by social media platforms. *Telecommunications Policy, 39*(9), 761–770. https://doi.org/10.1016/j.telpol.2015.04.003

Durand, A., Hamersma, M., Van Oort, N., & Hoogendoorn, S. (2024). Digital engagement for travel information among car and public transport users in the Netherlands. *Transportation Research Interdisciplinary Perspectives, 28*, Article 101285. https://doi.org/10.1016/j.trip.2024.101285

Durand, A., Zijlstra, T., Hamersma, M., 'T Hoen, A., Van Oort, N., Hoogendoorn, S., & Hoogendoorn-Lanser, S. (2023). Fostering an inclusive public transport system in the digital era: An interdisciplinary approach. *Transportation Research Interdisciplinary Perspectives, 22*, Article 100968. https://doi.org/10.1016/j.trip.2023.100968

Durand, A., Zijlstra, T., Hamersma, M., Hoen, A., Van Oort, N., Hoogendoorn-Lanser, S., & Hoogendoorn, S. (2023). "Who can I ask for help?": Mechanisms behind digital inequality in public transport. *Cities, 137*, Article 104335. https://doi.org/10.1016/j.cities.2023.104335

Durand, A., Zijlstra, T., Van Oort, N., Hoogendoorn-Lanser, S., & Hoogendoorn, S. (2022). Access denied? Digital inequality in transport services. *Transport Reviews, 42*(1), 32–57. https://doi.org/10.1080/01441647.2021.1923584

Edwards, M., & Mackett, R. L. (1996). Developing new urban public transport systems. *Transport Policy, 3*(4), 225–239. https://doi.org/10.1016/S0967-070X(96)00023-6

Faith, B., Hernandez, K., & Beecher, J. (2022). *Digital poverty in the UK*. Institute of Development Studies. https://doi.org/10.19088/IDS.2022.057

Feuerriegel, S., Maarouf, A., Bär, D., Geissler, D., Schweisthal, J., Pröllochs, N., … Van Bavel, J. J. (2025). Using natural language processing to analyse text data in behavioural science. *Nature Reviews Psychology, 4*(2), 96–111. https://doi.org/10.1038/s44159-024-00392-z

Frederico, C. S., Pereira, A. L. S., Marte, C. L., & Yoshioka, L. R. (2021). Mobile application for bus operations controlled by passengers: A user experience design project (UX). *Case Studies on Transport Policy, 9*(1), 172–180. https://doi.org/10.1016/j.cstp.2020.11.014

Garritsen, K.É., Grigolon, A. B., & Geurs, K. T. (2025). The role of digital mobility skills in the uptake of shared modes at mobility hubs. *Journal of Urban Mobility, 7*, Article 100111. https://doi.org/10.1016/j.urbmob.2025.100111

Gloutnikov, S. S. (2018). *Using sentiment analysis and pattern matching to signal user review abnormalities*. Master of Science. San Jose State University. https://doi.org/10.31979/etd.8hau-p2f5.

Graham, H., De Bell, S., Flemming, K., Sowden, A., White, P., & Wright, K. (2018). The experiences of everyday travel for older people in rural areas: A systematic review of UK qualitative studies. *Journal of Transport & Health, 11*, 141–152. https://doi.org/10.1016/j.jth.2018.10.007

Guzman, L. A., & Bocarejo, J. P. (2017). Urban form and spatial urban equity in Bogota, Colombia. *Transportation Research Procedia, 25*, 4491–4506. https://doi.org/10.1016/j.trpro.2017.05.345

Hargittai, E. (2003). *The digital divide and what to do about it*.

Heeks, R. (2021). From digital divide to digital justice in the global south: Conceptualising adverse digital incorporation. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3907633

Helsper, E. (2021). *The digital disconnect: The social causes and consequences of digital inequalities*.

Hirschberg, J., & Manning, C. D. (2015). Advances in natural language processing. *Science, 349*(6245), 261–266.

Hutto, C., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media, 8*(1), 216–225. https://doi.org/10.1609/icwsm.v8i1.14550

Ji, T., Chen, J.-H., Wei, H.-H., & Su, Y.-C. (2021). Towards people-centric smart city development: Investigating the citizens' preferences and perceptions about smart-city services in Taiwan. *Sustainable Cities and Society, 67*, Article 102691. https://doi.org/10.1016/j.scs.2020.102691

Jin, H., Wang, Q., Yang, Y.-F., Zhang, H., Gao, M.(M.), Jin, S., … Chuan-Peng, H. (2023). The Chinese Open Science Network (COSN): Building an Open Science community from scratch. *Advances in Methods and Practices in Psychological Science, 6*(1), Article 25152459221144986. https://doi.org/10.1177/25152459221144986

Kopackova, H., & Komarkova, J. (2020). Participatory technologies in smart cities: What citizens want and how to ask them. *Telematics and Informatics, 47*, Article 101325. https://doi.org/10.1016/j.tele.2019.101325

Liu, W., Xu, C. K., & Zhang, Z. (2025). The digital divide in shared bicycle use: Does smart mobility technology reproduce transport inequality? *Cities, 158*, Article 105706. https://doi.org/10.1016/j.cities.2025.105706

Liu, X., Faes, L., Kale, A. U., Wagner, S. K., Fu, D. J., Bruynseels, A., , … Kern, C., et al. (2019). A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: A systematic review and meta-analysis. *The Lancet Digital Health, 1*(6), e271–e297.

Lucas, K. (2012). Transport and social exclusion: Where are we now? *Transport Policy, 20*, 105–113.

Ma, Z., Hu, G., Lin, T.-S., Li, L., Hu, S., Hagen, L., & Baecher, G. B. (2025). Analyzing public response to wildfires: A socio-spatial study using SIR models and NLP techniques. *Computers, Environment and Urban Systems, 121*, Article 102333. https://doi.org/10.1016/j.compenvurbsys.2025.102333

Mahmud, M. S., Jaman Bonny, A., Saha, U., Jahan, M., Tuna, Z. F., & Al Marouf, A. (2022). Sentiment analysis from user-generated reviews of ride-sharing mobile applications. In *2022 6th International Conference on computing methodologies and communication (ICCMC)* (pp. 738–744). https://doi.org/10.1109/ICCMC53470.2022.9753947

Manandhar, B., Dunkel Vance, K., Rawat, D. B., & Yilmaz, N. (2025). Leveraging digital twin technology for sustainable and efficient public transportation. *Applied Sciences, 15*(6).

Martin-Sanromán, J.-R., Carballo, F. S., Rubio, F. G., & Raposo, D. (2019). Aesthetics, usability, and the digital divide: An approach to beauty, UX, and subjective well-being. In *The Routledge Handbook of positive communication* (pp. 286–294). Routledge.

Nguyen, M. H., & Hargittai, E. (2023). Digital disconnection, digital inequality, and subjective well-being: A mobile experience sampling study. *Journal of Computer-Mediated Communication, 29*(1), Article zmad044. https://doi.org/10.1093/jcmc/zmad044

Nie, Y., Bhatnagar, S., Smith, D., & Suel, E. (2024). Disparities in public transport accessibility in London from 2011 to 2021. *Computers, Environment and Urban Systems, 113*, Article 102169. https://doi.org/10.1016/j.compenvurbsys.2024.102169

Nurminen, V., Rossi, S., Rinne, T., & Kyttä, M. (2024). How has digital participatory mapping influenced urban planning: Views from nine planning cases from Finland. *Computers, Environment and Urban Systems, 112*, Article 102152. https://doi.org/10.1016/j.compenvurbsys.2024.102152

Ojadi, J. O., Onukwulu, E. C., Odionu, C. S., & Owulade, O. A. (2023). Natural language processing for climate change policy analysis and public sentiment prediction: A data-driven approach to sustainable decision-making. *7*(3).

Owen, D., Hogarth, T., & Green, A. E. (2012). Skills, transport and economic development: Evidence from a rural area in England. *Journal of Transport Geography, 21*, 80–92. https://doi.org/10.1016/j.jtrangeo.2012.01.015

Pineda-Jaramillo, J., Fazio, M., Pira, M. L., Giuffrida, N., Inturri, G., Viti, F., & Ignaccolo, M. (2023). A sentiment analysis approach to investigate tourist satisfaction towards transport systems: The case of Mount Etna. *Transportation Research Procedia, 69*, 400–407. https://doi.org/10.1016/j.trpro.2023.02.188

Richardson, L. (2007). *Beautiful Soup Documentation*.

Rinker, T. (2016). *sentimentr: Calculate Text Polarity Sentiment (p. 2.9.0) [Data set]*. https://doi.org/10.32614/CRAN.package.sentimentr

Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., … Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science, 58*(4), 1064–1082. https://doi.org/10.1111/ajps.12103

Ruiu, M. L., Ruiu, G., & Ragnedda, M. (2025). Beyond access: Motivation and digital literacy in sustainable technology use. *International Journal of Communication, 19*, 25.

Shekarappa G., S., Badi, M., Raj, S., & Mahapatra, S. (2023). An overview of smart city planning—The future technology. In *Artificial intelligence and machine learning in Smart City planning* (pp. 319–335). Elsevier. https://doi.org/10.1016/B978-0-323-99503-0.00016-8.

Shelton, T., Zook, M., & Wiig, A. (2015). The 'actually existing smart city'. *Cambridge Journal of Regions, Economy and Society, 8*(1), 13–25. https://doi.org/10.1093/cjres/rsu026

Shtebunaev, S., Gullino, S., & Larkham, P. J. (2023). Planning the smart city with young people: Teenagers' perceptions, values and visions of smartness. *Urban Planning, 8*(2). https://doi.org/10.17645/up.v8i2.6411

Sohail, A., & Zhang, L. (2025). Using large language models to facilitate academic work in the psychological sciences. *Current Psychology, 44*(9), 7910–7918. https://doi.org/10.1007/s12144-025-07438-2

Van Deursen, A. J. A. M., & Helsper, E. J. (2015). The third-level digital divide: Who benefits Most from being online? In L. Robinson, S. R. Cotten, J. Schulz, T. M. Hale, & A. Williams (Eds.), *Vol. 10. Studies in media and communications* (pp. 29–52). Emerald Group Publishing Limited. https://doi.org/10.1108/S2050-206020150000010002.

Van Dijk, J. A. G. M. (2006). Digital divide research, achievements and shortcomings. *Poetics, 34*(4–5), 221–235. https://doi.org/10.1016/j.poetic.2006.05.004

Verlinghieri, E., & Schwanen, T. (2020). Transport and mobility justice: Evolving discussions. *Journal of Transport Geography, 87*, Article 102798. https://doi.org/10.1016/j.jtrangeo.2020.102798

Verma, S. (2022). Sentiment analysis of public services for smart society: Literature review and future research directions. *Government Information Quarterly, 39*(3), Article 101708. https://doi.org/10.1016/j.giq.2022.101708

Vicente-Saez, R., & Martinez-Fuentes, C. (2018). Open Science now: A systematic literature review for an integrated definition. *Journal of Business Research, 88*, 428–436.

Wilson, A., Tewdwr-Jones, M., & Comber, R. (2019). Urban planning, public participation and digital technology: App development as a method of generating citizen involvement in local planning processes. *Environment and Planning B: Urban Analytics and City Science, 46*(2), 286–302. https://doi.org/10.1177/2399808317712515

Yan, X., Guo, J., Lan, Y., & Cheng, X. (2013). A biterm topic model for short texts. In *Proceedings of the 22nd International Conference on World Wide Web* (pp. 1445–1456).

Yap, W., Janssen, P., & Biljecki, F. (2022). Free and open source urbanism: Software for urban planning practice. *Computers, Environment and Urban Systems, 96*, Article 101825.

Zukeran, D. P., Carter, C. E., & Triviño, M. H. (2024). The political dimension of making a place: Framing the right to the City in placemaking. In *Placemaking: People, properties, planning* (pp. 163–177). Emerald Publishing Limited.