Social Media Metrics for Urban Mobility: Engagement Insights from X (formerly Twitter) in European Cities

Maria Karatsoli[1*], Ioanna Pagoni[1], Kimon Deligiannis[2], Stefanos Vlachos[2], Christos Tryfonopoulos[2], Paraskevi Raftopoulou[2], Amalia Polydoropoulou[1]

1. Department of Shipping, Trade and Transport, University of the Aegean, Chios, Greece
2. Department of Informatics and Telecommunications, University of the Peloponnese, Tripoli, Greece

*mkaratsoli@aegean.gr

Abstract

Social media has become a key tool for urban mobility communication, allowing municipalities and transport operators to share updates, interact with the public, and monitor sentiment. This study explores how seven European cities use X (formerly Twitter) to engage on transport-related issues. Data from 15 official X accounts were filtered using mobility-related keywords to include only relevant posts. Three indices were developed to quantify performance: the User Engagement Metric Index (UMEI), the Transport Content Activity Index (TCAI), and the Public Sentiment Alignment Score (PSAS). Each index was standardized and scaled to enable meaningful comparison across accounts and cities. The results show significant variation among cities, which was further examined in relation to population size, complexity of public transport systems, and national rates of social media use. The findings offer insights into how city characteristics influence digital engagement and provide a structured approach to improve social media strategies in urban mobility planning.

Keywords: *Sentiment analysis, Public Engagement, Sustainable Mobility, Data Analytics, Hashtags, Digital Strategies, Smart Mobility, Communication*

1. *Introduction*

Promoting sustainable urban transport is important for addressing the environmental, social, and economic challenges faced by current cities. The transition to more sustainable mobility patterns, particularly through the increased use of public transport, active modes, and shared mobility, plays a key role in achieving climate-neutral cities, reducing greenhouse gas emissions, and managing the negative impacts of road transport such as congestion, air pollution, and noise. However, to achieve this, it is critical to understand how citizens and public authorities interpret, engage with and interact with transport systems, while also ensuring that these interactions lead to improved user satisfaction and overall service effectiveness.

Traditionally, public perception of mobility services has been assessed through large-scale surveys, interviews, and focus groups. While these methods are valuable, they are also costly, time-consuming, and limited in scope and frequency while often capturing a static picture rather than ongoing public sentiment, making it difficult for authorities to respond quickly to shifting attitudes or emerging issues.

In recent years, social media has emerged as a valuable source of public opinion and behavioral insight in urban contexts. Platforms like X and Facebook allow users to spontaneously express views, experiences, and frustrations related to transport services in real time. Compared to traditional methods, social media offers large volumes of unprompted feedback that can be mined continuously and at relatively low cost. Studies have suggested that social media posts may provide more immediate expressions of user satisfaction or dissatisfaction, especially when users perceive these platforms as informal and accessible channels for public (Balbi et al., 2018; Yang et al., 2020; Nikolaidou et al., 2024).

In the field of transport research, social media has been increasingly used to monitor mobility patterns, track disruptions, and assess public sentiment toward specific policies or services (e.g., Cottrill et al., 2017; Zayet et al., 2021; Mathews et al., 2022). These applications highlight the potential of social media to complement traditional methods and provide richer, more dynamic understandings of public attitudes toward mobility systems. Yet, most existing studies either focus on specific events (e.g., strikes, delays, or emergencies) or analyze data from a single city, limiting generalizability.

This study addresses that gap by introducing a structured and quantifiable method to assess social media performance related to mobility communication across multiple European cities. By analyzing X activity from official accounts managed by municipalities and transport operators, the study develops and applies three composite indices to evaluate user engagement, content activity, and sentiment. These indices allow for cross-city comparisons and are interpreted considering contextual factors such as city size, public transport system complexity, and social media adoption rates. This approach not only provides insights into how cities perform digitally in the mobility domain but also offers a methodological blueprint for evaluating and improving online communication strategies aimed at promoting sustainable urban transport.

The structure of the paper is as follows: Section 2 reviews the background and related literature. Section 3 presents the methodological approach, covering data collection and the development of performance indices. Section 4 discusses the results, including index-based comparisons across cities and their relationship to city characteristics. Section 5 concludes with key findings and implications for urban transport communication.

2. Background

In recent years, social media has emerged as a pivotal platform for urban mobility communication, enabling municipalities and transport operators to disseminate information, engage with the public, and gather feedback. The interactive nature of social media platforms allows for real-time communication, fostering a two-way dialogue between service providers and users, which is essential for promoting sustainable transport options and enhancing user satisfaction.

Studies have highlighted the significant impact of social media communication on transportation behavior. For instance, increased social media activity correlates with higher usage of public and sustainable transportation modes, such as subways and bicycle-sharing systems, while private car usage tends to decrease (Martinez-de-Ibarreta et al., 2024). This suggests that effective social media strategies can influence public transport adoption and support environmental goals.
Moreover, social media platforms serve as valuable tools for collecting unprompted user feedback, offering insights into public sentiment and experiences. This user-generated content can inform service improvements and policy decisions, providing a cost-effective alternative to traditional survey methods (Karatsoli & Nathanail, 2023).

The integration of social media analytics into urban mobility planning allows for a more responsive and adaptive approach to transport management. By leveraging data from platforms like X, authorities can

monitor public reactions to service changes, identify areas of concern, and implement timely interventions to address issues.
In summary, social media plays a crucial role in urban mobility communication by facilitating information dissemination, enhancing public engagement, and providing actionable insights through user feedback (Karatsoli & Nathanail, 2021). Its strategic use can contribute to more sustainable and user-centered urban transport systems.

The growth of social media, accommodating billions of users, has prompted various stakeholders in the transport and mobility sector to develop social media mining methodologies, such as public opinion mapping and research, with the aim of extracting valuable insights and data-driven decisions. In this context, Carvalho et al. (2021) presents a cross-platform data harvesting process, ingesting data from both X and Instagram to produce user mobility patterns within the city of Porto in Portugal, while Kuflik et al. (2017) exploit X data by gathering user transport-oriented Tweets, over a three-day period before and after three Liverpool football matches held during 2012. However, both papers leverage X's API, which comes with a number of restrictions including its pricing, the time range of access to Tweets, and the level of usage based on the API subscription plan. Contrary, developing a web scraping procedure makes data collection more versatile, even if it suffers from occasional website changes (Dongo et al., 2020). From the various web scraping practices aligned with the transport and mobility content encountered across X, the solutions introduced by Balla et al. (2023); Qi et al., (2020); Sánchez-Ávila et al. (2020), demonstrate similar strategies: they initially crawl X to create a dataset consisting of user posts, followed by Natural Language Processing (NLP) procedures, capable of extracting valuable knowledge. Specifically, Balla et al. (2023), measure public sentiment with respect to the use of electric vehicles in a dataset assembled from X between 2012-2022, with the aim of deriving improved guidelines for electric transportation. Similarly, Qi et al. (2020) present a public opinion mining method on transportation services of Miami-Dade County, deploying Post Intensity and Average Sentiment Analysis metrics on Twitter data, which were harvested by implementing a web crawler. Sánchez-Ávila et al., 2020 provide a framework that adopts the T-Hoarded tool to capture X content corresponding to terms related to incidents and barriers involved in pedestrian mobility, which have been defined in advance, while deploying Named Entity Recognition (NER) to detect location objects in users' trajectories.

A user-friendly, zero-cost, collaborative datalake, capable of launching social media (Facebook, X and TripAdvisor) crawlers, storing and manipulating collected social media content in tailored designed data repositories, while providing straightforward built-in filtering and analysis tools (including NLP and NER) to transform aggregated information into insightful charts and analytics, is provided by Deligiannis et al. (2020). Along the same lines, approaches such as Hirata & Matsuda (2023), Tran et al. (2023) and Chen et al. (2023), make use of NLP tools to perform Sentiment Analysis to transportation data within the COVID-19 pandemic. Hirata & Matsuda (2023) demonstrate a Sentiment Analysis process employing the BERT Large Language Model (LLM) on X logistics-based content in Japan. Tran et al. (2023) introduce a Machine Learning (ML-based Sentiment Analysis methodology, aiming to enhance conventional transportation data acquisition strategies, by utilizing a partitioned X corpus that consists of the travel experiences of approximately 120K individuals using the Vancouver Metro in Canada prior to and during the pandemic. In a similar way, the solution presented by Chen et al. (2023) exploits content from X to formulate travel type classifiers, while Sentiment Analysis is implemented to comprehend people's transportation type preference within the pandemic period in New York City. Finally, both solutions reported by Fontes et al. (2023) and Stiebe (2024) put forward equivalent twofold procedures that involve travel-related filtering on the data collected in the first stage, while as a second step Sentiment Analysis is performed on the refined outcome to extract fundamental insights that shape transportation decision-making. More precisely, the work by Fontes et al. (2023), proposes an NLP approach applied to data extracted from X, which takes advantage of the pre-trained BERT LLM for the purpose of categorizing Tweets associated with transportation in the cities of London, Melbourne and New York. Stiebe (2024) conducts a multi-platform social media analysis focusing on sustainable mobility and transport, aiming to improve the awareness of socio-technical transitions to low-carbon transport, by initiating the web scrapers Twint and Instaloader with predefined hashtag terms to capture X and Instagram content respectively, while also performing quantitative text measurements and Sentiment Analysis on user posts and image captions.

3. *Methodological approach*

This section outlines the approach followed to collect, process, and analyze transport-related X data from the seven European cities (Antwerp, Bologna, Gdynia, Las Palmas, Rouen, Tallinn, Valladollid). These cities share different characteristics in terms of mobility, demographics and user engagement in social media. The aim was to convert the unstructured collected data into structured insights that can support evidence-based decision-making in urban transport communication and public engagement strategies. The analysis focused on X activity from accounts linked to municipalities and public transport operators in nine European cities.

3.1 Data collection and preparation

To analyze public engagement and sentiment with urban mobility content on social media, a dataset was compiled based on 15 X accounts linked to municipalities (7 accounts) and public transport operators (8 accounts) in seven European cities. These were selected from an initial set of 18 accounts, based on a keyword-based filtering process to ensure that only transport-related content was retained. Data was collected using a custom scraping pipeline that extracted both regular and retweeted content, ensuring coverage of user generated as well as shared communication.

Each tweet was associated with several metadata insights, including the number of likes, retweets, views, and replies, along with sentiment polarity and subjectivity scores (Matthew & Varghese, 2023) calculated through natural language processing tools (Vasiliev, 2020). The collected dataset also included account-level metadata such as the number of followers, total tweets, and inactivity periods. All metrics were carefully organized to enable normalization, aggregation, and multi-dimensional interpretation across accounts and cities.

As mentioned earlier, social media platforms, including X, host a vast amount of user-generated content that reflects people's preferences and experiences. In our case, accounts of transportation companies and municipalities were identified as valuable sources of information for assessing users' transportation experiences. To this end, an additional pipeline for collecting and analyzing data from X was developed. This pipeline consists of 3 main steps (see Figure 2): (i) data collection: automated collection of tweets from the target accounts, (ii) data preparation: cleaning and preparation of data for analysis, and (iii) data analysis and visualization. The following sections describe each stage in detail.

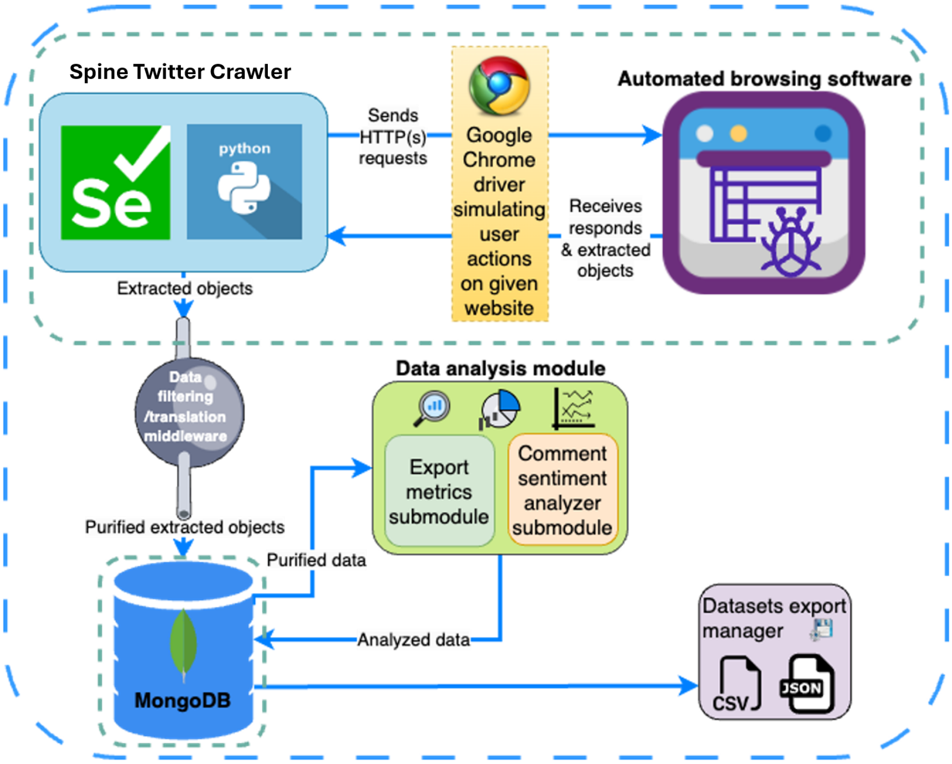

***Figure 1:** The X data collection and analysis system architecture.*

Data Collection: X is a platform with hundreds of millions of active users generating content daily. This calls for an automated method of collecting data, since manual effort would be a labor-intensive and tedious procedure. The large volume of data has also affected the adopted strategies for rendering content on screen, with a large share of accessible data being dynamically loaded to promote scalability. These requirements, in combination with the time and resource restrictions coming with the use of the official X API, highlighted the Selenium framework and Python as the best candidates for this data collection task. In more detail, Selenium was configured to log in the platform via the Chrome browser, visit 18 X accounts of municipalities and transportation companies, scroll throughout the uploaded content and collect tweets, as well as their comments, that had been uploaded within the time frame 01/01/2022 – 30/06/2023. For each account, the crawler collected the account name, description, national origin, links to external sources, date of joining X, number of followers, and a list of target tweet URLs. Similarly, for each tweet, the upload date, tweet text (including emojis, mentions, and hashtags), an indicator of whether it was an original post or a retweet, and the number of likes, retweets and views was gathered. After iteratively loading the replies, the crawler also extracted the text and corresponding engagement metrics (likes, retweets and views) for each reply. Collected data were stored in the form of documents in a MongoDB database. The data collection process resulted in 7.914 tweets, which can be exported and distributed via JSON or CSV files. Finally, it is worth highlighting that the data was collected solely for academic research purposes and exclusively publicly available content was targeted, with no intent to infringe on user privacy or exploit the platform commercially. Efforts were made to minimize impact on the platform and to ensure compliance with ethical research standards, while no personally identifiable information violating GDPR regulation was either collected or stored in the database.

Data Preparation: The goal of the data collection task was to gather and analyze user-generated replies to transportation-related tweets. However, since the selected X accounts span across different European countries, users' content may be produced in various languages. Therefore, it was deemed necessary to translate all tweets and replies into English using the Deep Translator Python framework

(Baccouri, 2020). This framework parsed the text, detected the source language with a certain level of confidence and translated it into English, excluding non-text content (e.g., emojis) from the translation process. In addition to overcoming the language barrier, the analysis should focus solely on tweets related to transportation. This was achieved by filtering out tweets that contained no transportation-related keywords (e.g., public transportation, traffic, mobility).

Data Analysis and Visualization: The objective of assessing users' experiences and opinions based on the collected social media data was approached from three different but complementary directions: (i) analyzing the activity of the target accounts, (ii) measuring community engagement with the uploaded content, and (iii) evaluating the sentiment expressed in user replies. Regarding activity, metrics such as the average number of daily tweets, the average interval between consecutive tweets and the ratio of self-generated tweets to retweets were calculated. Community engagement was assessed by analyzing the number of reactions (likes, views, retweets) and replies, both in total and on average per tweet, while sentiment analysis, employed the SpaCy NLP ecosystem. This framework applied rule-based sentiment analysis to the textual data and provided four outputs: (i) Polarity: a decimal between -1 and 1 indicating sentiment from negative to positive; (ii) Subjectivity: a decimal between 0 and 1 reflecting the degree of personal opinion versus objectivity; (iii) Sentiment Assessments: a list of detected terms along with their associated polarity and subjectivity scores; and (iv) trigrams: sequences of adjacent words. The polarity and subjectivity metrics were used to assess the emotional state of tweet replies and, in combination with the measurements, contributed to acquiring an account-specific estimation of users' response to transportation-related content.

3.2 Index development

Three indices were developed to convert a large and diverse set of engagement and sentiment metrics into interpretable results. These indices serve as quantitative representations of different dimensions of X-based transport communication: public engagement, content activity, and alignment with user sentiment.

The first index, **User Engagement Metric Index (UMEI)**, quantifies the degree of interaction a post generates from users. It aggregates the average number of likes, retweets, replies, and views for each account, as these are key indicators of how well transport-related content reaches and affects the audiences. Each of these four metrics is converted into a standardized z-score and then rescaled from 0 to 10.

$$UMEI = \frac{1}{4} \times \left(z_{likes} + z_{retweets} + z_{replies} + z_{views}\right) \text{ (1)}$$

The second index, **Transport Content Activity Index (TCAI)**, measures the proportion of tweets that include transport-related keywords. It is designed to reflect the degree of thematic focus that the account places on transport issues, highlighting whether communication efforts are aligned with mobility topics.

$$TCAI = \frac{1}{2} \times \left(z_{keyword_{proportion}} + z_{tweets_{per_{day}}}\right) \text{ (2)}$$

The third index, **Public Sentiment Alignment Score (PSAS)**, evaluates the overall tone and subjectivity of the transport-related tweets. Higher polarity values reflect more positive perceptions, while subjectivity indicates whether the tweets represent facts or personal opinions. This index captures how public reactions align with the messaging tone of the account.

$$PSAS = \frac{1}{2} \times (z_{polarity} + z_{subjectivity}) \text{ (3)}$$

To enable cross-variable comparability, each z-score is rescaled to a [0, 10] range using the following transformation:

$$Score_{scaled} = \frac{z - z_{min}}{z_{max} - z_{min}} \times 10 \quad (4)$$

This approach avoids any bias by the different metric scales and allows all index components to contribute equally. The resulting indices allow direct interpretation, where higher scores indicate stronger engagement, higher content activity, or more positive public sentiment in relation to urban transport.

4. Results and City-Level Comparisons

This section presents the main findings of the analysis across the selected European cities. It highlights key differences in social media performance and explores possible factors that help explain these variations. The results are structured to allow both individual account evaluation and city-level comparison.

4.1 Index-Based Assessment of Social Media Performance

The three transport-related indexes, UMEI (User Mobility Engagement Index), TCAI (Transport Communication Activity Index), and PSAS (Public Sentiment Alignment Score), were calculated for each one of the X accounts. These indexes were then aggregated for each one of the seven selected European cities, covering metrics such as average likes, retweets, replies, views, tweet frequency, and sentiment.

Table 1 summarizes the normalized index values by X accounts (M: Municipality account; PT: Public Transport operator) and the aggregated scores per selected city. In terms of user engagement (UMEI), Ayuntamiento de Valladolid (PT) records the highest score, reflecting strong audience interaction with its content. Other accounts such as Guaguas Municipales (M) and Comune di Bologna (M) also show relatively high engagement levels, indicating frequent user reactions in the form of likes, replies, and retweets. Conversely, accounts such as Réseau Astuce (PT) and Marconi Express (PT) register low UMEI values, pointing to limited user response despite their presence on the platform.

When examining content focus (TCAI), Réseau Astuce (PT), De Lijn (PT), and NMBS (PT) exhibit the highest levels of transport-related activity. These accounts post frequently and maintain strong alignment with mobility-related topics, as reflected in their keyword density and tweet frequency. In contrast, Ayuntamiento de Valladolid (PT) and Gdynia (M) demonstrate minimal transport focus, suggesting that their social media strategies are either more general or less consistently related to mobility themes.

In terms of sentiment (PSAS), Gdynia stands out with the highest value, indicating positive and emotionally related engagement with users. Slim naar Antwerpen (M) and Ayuntamiento de Las Palmas (M) also reflect strong sentiment alignment. Lower values in accounts such as Marconi Express (PT) and CittàmetroBologna (PT) suggest a more neutral or emotionally disconnected public response.

Table 1: *Index scores of the X accounts and the seven European cities*

Account	UMEI	TCAI	PSAS	Aggregated UMEI	Aggregated TCAI	Aggregated PSAS	City
De Lijn (PT)	1,39	8,208	4,339	1,55	6,71	5,46	Antwerp
NMBS (PT)	1,50	8,200	5,655				
Slim naar Antwerpen (M)	1,75	3,735	6,383				
CittàmetroBologna (PT)	1,00	1,454	3,171	1,77	1,48	3,53	Bologna
Comune di Bologna (M)	3,89	0,976	5,038				
Marconi Express (PT)	0,43	2,011	2,384				
Gdynia (M)	0,79	0,162	**9,296**	0,79	0,162	9,296	Gdynia
Ayuntamiento de Las Palmas de Gran Canaria (M)	2,52	0,483	7,046	2,28	2,97	5,76	Las Palmas
Guaguas Municipales (PT)	2,83	3,553	5,154				
Guaguas Global (PT)	1,48	4,864	5,095				
Métropole Rouen Normandie (M)	1,05	1,486	3,753	0,71	5,74	4,42	Rouen
Réseau Astuce (PT)	0,38	**10,000**	5,092				
Tallinn-European Green Capital 2023 (M)	1,15	2,006	3,459	1,15	2,01	3,46	Tallinn
AuvasaVLL (M)	1,30	3,652	6,224	4,52	1,85	6,01	Valladollid
Ayuntamiento de Valladolid (PT)	**7,74**	0,053	5,800				

At the city level, the aggregated indices allow for a comparative view of how different urban contexts perform across the three key dimensions of social media activity. The following radar chart (Figure 1), which summarizes these scores for the seven cities, reveals clear performance patterns. Antwerp presents a well-rounded profile, with high TCAI and solid engagement and sentiment scores, indicating a strategy that balances content volume and relevance with audience responsiveness. Las Palmas also performs strongly, particularly in UMEI and PSAS, suggesting that its communication affects emotionally users and generates active engagement, even if its transport content ratio is slightly lower than top performers. Bologna shows strong UMEI and PSAS, highlighting its ability to maintain emotional connection and interaction, despite limited transport-specific content, which may indicate broader municipal priorities on X. On the other hand, Rouen and Tallinn maintain moderate values across all three indices, which might point to consistent but less intensive strategies, while Valladolid presents a unique profile, with very high engagement but minimal transport-related content. Gdynia, by contrast, shows limited content and engagement scores, yet excels in sentiment alignment, possibly reflecting the tone or emotional impact of the few posts it publishes.

This index-based view emphasizes the diversity in communication strategies and public responses across cities. While some cities achieve balance, others excel in one or two dimensions, underlining the value of a multidimensional assessment approach. These differences prompt further investigation into underlying factors such as the city's population size or the complexity and complexity of its public transport system, that may shape digital communication performance.

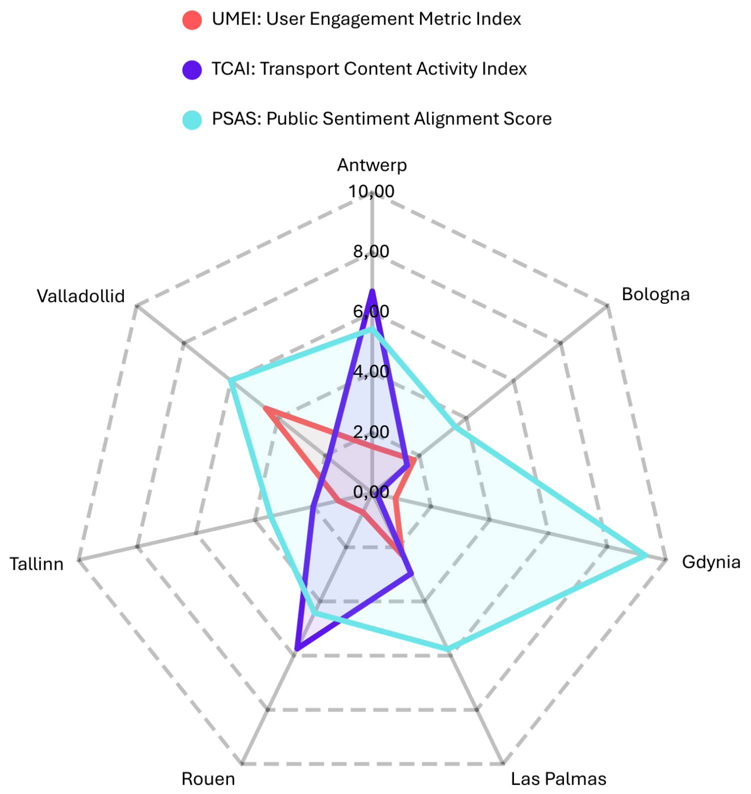

***Figure 2:** The three indexes across seven European cities*

The next section explores these aspects in more detail, aiming to contextualize the observed index scores by relating them to key city characteristics. This additional analysis helps to clarify whether strong social media performance is a function of communication effort alone, or if it also reflects broader urban and mobility system conditions.

4.2 City Characteristics and Their Impact on Social Media Scores

To explore potential factors behind the observed differences three contextual factors were considered: population size, the complexity of the public transport (PT) system, and social media usage levels at the national scale (Table 2). **Population size** was selected as a representative metric of urban density and potential audience reach. Data from official statistical sources revealed that the cities in the study vary significantly in size—from Rouen with approximately 113,000 residents (INSEE, 2021) to Antwerp with around 540,000 (Statbel, 2023). These differences could have an impact on the expected level of public engagement, and the communication needs of local authorities and transport operators. **Public transport complexity** was evaluated based on the availability of multimodal options such as bus, tram, metro, and rail. Cities like Antwerp, Bologna, Gdynia, and Tallinn offer integrated systems combining multiple modes (Multimodal -Advanced), while Las Palmas, Rouen, and Valladolid primarily rely on bus networks (Bus based- Basic). The classification was derived from each city's public transport portals and verified through local mobility reports (e.g., Tallinna Transpordiamet, 2023; Bologna Urban Transport Plan, 2021). To account for the broader digital environment in which these X accounts operate, **national-level social media usage** was included. These percentages, representing the proportion of internet users who actively use social media, were obtained from DataReportal's 2024 Digital Global Overview. Estonia, for instance, showed a relatively high social media penetration rate of 75%, while France and

Italy had lower levels around 58%, possibly affecting the base level of user engagement expected in those contexts.

These three characteristics were compared against the aggregated social media index scores (UMEI, TCAI, PSAS) to identify potential patterns (Table 2). It was observed that cities with larger populations and more advanced PT systems, such as Antwerp and Bologna, tended to have higher overall user engagement and more consistent activity levels. Conversely, cities like Rouen, which combine a smaller population with a bus-only system, scored lower on the UMEI, indicating limited reach and interaction.

Interestingly, a stronger presence on social media did not always correspond with city's PT system. Valladolid, which has a less advanced PT network, showed high UMEI scores due to a very active and responsive X presence. This suggests that while structural factors influence potential reach, content strategy and posting frequency remain decisive.

Overall, the findings indicate that both structural context and communication practices play a role in shaping social media performance. Cities with complex transport networks, large populations, and high national digital engagement may have a greater incentive and potential to build dynamic and interactive online platforms. However, targeted content and regular activity can raise the presence of even smaller, less-connected cities.

Table 2: *Index scores of the X accounts and the seven European cities*

City	Population (approx..)	PT compexity	Social media usage (%internet users – country level)	UMEI	TCAI	PSAS
Antwerp	540.000	Advanced	67% (Belgium)	1,55	6,71	5,46
Bologna	390.000	Advanced	58% (Italy)	1,77	1,48	3,53
Gdynia	246.000	Advanced	71% (Poland)	0,79	0,162	9,296
Las Palmas	381.000	Basic	73% (Spain)	2,28	2,97	5,76
Rouen	113.000	Basic	58% (France)	0,71	5,74	4,42
Tallinn	440.000	Advanced	75% (Estonia)	1,15	2,01	3,46
Valladollid	298.000	Basic	73% (Spain)	4,52	1,85	6,01

5. *Conclusions*

This study introduced a structured framework for analyzing how municipalities and public transport operators in seven European cities use X to communicate urban mobility issues. Through the development of three indices, User Engagement Metric Index (UMEI), Transport Content Activity Index (TCAI), and Public Sentiment Alignment Score (PSAS), the research provided a consistent method to evaluate digital engagement, thematic focus, and alignment with public sentiment.

The results revealed considerable variation across cities. Higher scores were generally associated with accounts that posted more frequently, maintained a clear transport-related focus, and demonstrated stronger alignment with users' responses. This indicates that well-managed social media accounts can support more effective communication around mobility services and may reflect higher levels of public trust and interaction.

While the approach offers meaningful insights, certain limitations should be acknowledged. The analysis focused solely on X and on transport-related content identified through keyword filtering, which may not capture the full range of public hearing. Furthermore, sentiment was assessed using standard natural language processing tools, which, while efficient, may not accurately interpret complex or sarcastic expressions every time. The scope was also limited to a selected group of cities, and platform-specific features, such as algorithmic visibility, were not considered.

Despite these constraints, the methodology enables a systematic assessment of social media performance in the transport sector and can be adapted to other contexts. Future work could expand the platform scope and usability by supporting other social media sources, incorporate richer sentiment tools, or integrate real-world transport metrics to strengthen policy relevance. Additionally, longitudinal studies may help track changes in communication strategies over time, supporting public authorities and operators to improve their digital engagement practices.

6. *Acknowledgements*

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101096664.

7. *References*

Baccouri, N. (2020). Welcome to deep_translator's documentation! - deep_translator documentation. https://deep-translator.readthedocs.io/en/latest/index.html

Balbi, S., Misuraca, M., & Scepi, G. (2018). Combining different evaluation systems on social media for measuring user satisfaction. Information Processing & Management, 54(4), 674-685.

Balla, S. N., Pani, A., Sahu, P. K., & González-Feliu, J. (2023). Examining shifts in public discourse on electric mobility adoption through Twitter data. Transportation research part D: transport and environment, 121, 103843.

Carvalho, A. M., Ferreira, M. C., & Dias, T. G. (2021). Understanding mobility patterns and user activities from geo-tagged social networks data. *Transportation research procedia*, 52, 493-500.

Chen, X., Wang, Z., & Di, X. (2023). Sentiment analysis on multimodal transportation during the COVID-19 using social media data. Information 14 (2): 113.

Comune di Bologna. (2021). Piano Urbano della Mobilità Sostenibile (PUMS) – Bologna Metropolitan City. Retrieved from https://www.pumsbologna.it/

Cottrill, C., Gault, P., Yeboah, G., Nelson, J.D., Anable, J., & Budd, T. (2017). Tweeting Transit: An examination of social media strategies for transport information management during a large event. Transportation Research Part C: Emerging Technologies, Vol. 77, pp. 421–432.

Deligiannis, K., Raftopoulou, P., Tryfonopoulos, C., Platis, N., & Vassilakis, C. (2020). Hydria: An online data lake for multi-faceted analytics in the cultural heritage domain. Big Data and Cognitive Computing, 4(2), 7.

Dongo, I., Cadinale, Y., Aguilera, A., Martínez, F., Quintero, Y., & Barrios, S. (2020, November). Web scraping versus twitter API: a comparison for a credibility analysis. In Proceedings of the 22nd

International conference on information integration and web-based applications & services (pp. 263-273).

Fontes, T., Murços, F., Carneiro, E., Ribeiro, J., & Rossetti, R. J. (2023). Leveraging social media as a source of mobility intelligence: An NLP-based approach. IEEE Open Journal of Intelligent Transportation Systems, 4, 663-681.

Hirata, E., & Matsuda, T. (2023). Examining logistics developments in post-pandemic Japan through sentiment analysis of Twitter data. Asian Transport Studies, 9, 100110.

Institut National de la Statistique et des Études Économiques (INSEE). (2021). Populations légales 2019. Retrieved from https://www.insee.fr/fr/statistiques/5001880?sommaire=5001884

Karatsoli, M., & Nathanail, E. (2021). Social influence and impact of social media on users' mobility decisions. Journal of Sustainable Development of Transport and Logistics, 6(1), 32–48. https://doi.org/10.14254/jsdtl.2021.6-1.3

Karatsoli, M., Nathanail, E. (2023). Social Media and Urban Mobility Choices: How a Transport-Related Content Could Be Influential in Social Media. In: Nathanail, E.G., Gavanas, N., Adamos, G. (eds) Smart Energy for Smart Transport. CSUM 2022. Lecture Notes in Intelligent Transportation and Infrastructure. Springer, Cham. https://doi.org/10.1007/978-3-031-23721-8_67

Kemp, S. (2024). Digital 2024: Global Overview Report. DataReportal. Retrieved from https://datareportal.com/reports/digital-2024-global-overview-report

Kuflik, T., Minkov, E., Nocera, S., Grant-Muller, S., Gal-Tzur, A., & Shoor, I. (2017). Automating a framework to extract and analyse transport related social media content: The potential and the challenges. Transportation Research Part C: Emerging Technologies, 77, 275-291.

Martínez-de-Ibarreta, C., Cifuentes, J.A., Vallez, C.M. et al. Unveiling the effect of social media communication on urban mobility. Transportation (2024). https://doi.org/10.1007/s11116-024-10512-6

Mathews, S., Hou, J., Yigitcanlar, T., Golbabaei, F., Kankanamge, N., Downie, A., & Paz, A. (2025). From social media to policy: Public insights on transport communication and services in Queensland. Elsevier.

Matthew, A., & Varghese, L. (2023). Feature Based Sentiment Analysis Using Spacy & TextBlob. In Proceedings of the National Conference on Emerging Computer Applications (NCECA) (Vol. 5, No. 1, p. 445).

Nikolaidou, A., Georgiadis, G., Papaioannou, P., & Politis, I. (2024, September). Social Media as a Complementary Tool for Measuring Public Transport Customer Satisfaction. In Conference on Sustainable Urban Mobility (pp. 900-916). Cham: Springer Nature Switzerland.

Qi, B., Costin, A., & Jia, M. (2020). A framework with efficient extraction and analysis of Twitter data for evaluating public opinions on transportation services. Travel behaviour and society, 21, 10-23.

Sánchez-Ávila, M., Mouriño-García, M. A., Fisteus, J. A., & Sánchez-Fernández, L. (2020). Detection of barriers to mobility in the smart city using Twitter. IEEE Access, 8, 168429-168438.

Statbel – Statistics Belgium. (2023). Population on 1 January by region, province, district, and municipality. Retrieved from https://statbel.fgov.be/en/themes/population/structure-population

Stiebe, M. (2024). Social big data mining for the sustainable mobility and transport transition: findings from a large-scale cross-platform analysis. European Transport Research Review, 16(1), 28.

Tallinn Transport Department. (2023). Public transport in Tallinn. Retrieved from https://www.tallinn.ee/en/public-transport

Tran, M., Draeger, C., Wang, X., & Nikbakht, A. (2023). Monitoring the well-being of vulnerable transit riders using machine learning based sentiment analysis and social media: Lessons from COVID-19. Environment and Planning B: Urban Analytics and City Science, 50(1), 60-75.

Vasiliev, Y. (2020). Natural language processing with Python and spaCy: A practical introduction. No Starch Press.

Yang, F., Zhao, S., Li, W., Evans, R., & Zhang, W. (2020). Understanding user satisfaction with Chinese government social media platforms.

Zayet, T.M.A., Ismail, M.A., Varathan, K.D. et al. Investigating transportation research based on social media analysis: a systematic mapping review. Scientometrics 126, 6383–6421 (2021). https://doi.org/10.1007/s11192-021-04046-2