mobil.TUM 2024 – The Future of Mobility and Urban Space, April 10-12, 2024

The Voice of the City: Analysing User-Generated Content to Inform Active Mobility Policy Making Using Natural Language Models

Charlotte van Vessem [a*], Floriano Tori [a], Juliana Betancur Arenas [a], Vincent Ginis [a]

[a] *Vrije Universiteit Brussel, Belgium*

Keywords: *Urban mobility, Active mobility, Perceptions, Natural language models, User-generated content*

This work addresses the following topic(s) from the Call for Contributions:
(Please check at least one box)

☐ Placemaking to integrate urban spaces and mobility
☒ Promoting sustainable mobility choices in metropolitan regions
☐ Governing responsible mobility innovations
☒ Shaping the transition towards mobility justice
☒ System analysis, design, and evaluation
☐ other: _____

Extended Abstract

Problem statement

Transport planning relies heavily on traditional data collection methods such as manual surveys and traffic counts, yet these methods are expensive, time-consuming, and prone to errors. Moreover, these current data collection methods fail to consider multidimensional aspects of travel behaviour and the different experiences among their users' diversity. For instance, the survey's limitations include lower response rates and oftentimes homogenous samples, leading to the wrong assumption that the travel experiences of one group are universal, resulting in blind spots regarding characteristics such as gender, ethnicity, age, or income. In the context of the EU's ambition to become the first climate-neutral continent and with the deployment of plans and policies like the EU Green Deal and Fit for 55, there is a roadmap to shift towards more sustainable ways of moving, like active modes. However, there is a lack of understanding of citizens' perceptions and willingness to make part of this mobility shift. Emerging technologies like big data developments and advancing artificial intelligence can provide new methods and source material that were difficult (if not impossible) to utilise previously, complementing traditional research methods by extending the data scope. More specifically, pre-trained large language models, such as generative pre-trained transformers (GPT) or encoder models (Bert, XLM), provide an opportunity to gain an understanding of citizens' perceptions and feelings regarding changes in active mobility on a significantly larger scale than a manual analysis. These natural language models can be used to analyse user-generated content (UGC), such as social media, to help policymakers better understand mobility behaviour. Sentiment analysis based on UGC has the advantage that it is able to analyse populations that are otherwise difficult to reach with traditional data collection methods. This recent methodology and its scope have been underused in an urban mobility context, and only one analysis has been conducted on sentiment analysis of gender and mobility (Vasquez-Henriquez et al., 2020).

[*] *Corresponding author. Tel.: +32(0)26148329*
E-mail address: charlotte.van.vessem@vub.be

Research objectives

As part of the SSH CENTRE (Social Sciences and Humanities for Climate, Energy and Transport Research Excellence) collaborative call for transdisciplinary research between SSH and STEM, this research aims to explore the usability of natural language models for analysing UGC (such as tweets) to understand the public's perceptions regarding active and shared mobility as a whole and their related infrastructure. In order to consider a fair, inclusive, and just mobility transition, special attention will be paid to the differences in perceptions according to gender. Despite the differences in mobility patterns and habits between men and women, very little is known about the role of gender perceptions in the choice of transport modes. Tackling the data gender gap on people's perception towards these modes will help policymakers to accurately shape gender-inclusive mobility policy guidelines. This interdisciplinary and collaborative approach aims to expose a rigorous, replicable and valid methodology to collect qualitative data using some of the latest developments in technology and engineering to consider the gender gap in mobility and the relevance of the geographical and social context for the transition towards sustainable mobility.

Methodological approach

Given that X (formerly known as Twitter) is one of the most accessed mobile applications to report daily experiences, including transport-related ones, we decided to focus on this social network for the collection of user content. Data will be collected via posts on X discussing active and shared mobility in the Brussels Capital Region in Belgium (BCR) context. We consider posts (formerly known as tweets) between 2017-2022 located among the 19th communes of BCR, which included previously defined mobility keywords together with relevant stakeholders or relevant locations. The keywords consist of words and hashtags related to active and shared mobility words. The analysis will consider tweets in Dutch and French (as the two official languages of the region) and English (to account for tweets by the large population of non-Belgian natives). The next step will be to perform sentiment analysis on the collected posts using pre-trained large language models, classifying posts based on the sentiment expressed (i.e. positive, neutral or negative). The aim will be to understand public perception and feelings regarding the sustainable transition to active and shared modes of transport and infrastructure. Finally, we will explore if the users' responses can be classified according to their gender.

(Expected) results

Brussels has a fast-changing mobility context that has been quite controversial, aiming to increase the share of active modes in the city. Using this case study, we aim to develop a tool to understand better the nuances behind the resistance and support towards a sustainable mobility shift. Moreover, considering gender-sensitive data and analysis will contribute to our knowledge of gender-based differences, their influence on shared mobility usage and street transformation to encourage active mobility. This paper will respond to what extend variables mentioned in previous mobility studies such as gendered discourses and concerns related to street harassment and safety perception in the public space (Vasquez-Henriquez et al., 2020) can play a role in achieving a sustainable mobility shift. Moreover, it aims to identify gender-differentiated enablers and barriers to make the shift.

In general terms, this research project will be of value to policymakers by allowing them to understand the general public's perceptions in a broader way than they usually have access to by obtaining up-to-date qualitative data on their feelings and needs. We will also demonstrate whether methodologies based on large language models can be a complementary source of information during decision-making processes. While highlighting shortcomings in current methodologies to analyse citizens' perception of active mobility, our interdisciplinary approach will also provide a planning support tool that considers context-based and gender-based differences for better policy discussion and tailored strategies.

References

Vasquez-Henriquez P, Graells-Garrido E and Caro D (2020) Tweets on the Go: Gender Differences in Transport Perception and Its Discussion on Social Media. Sustainability 12(13). 13. Multidisciplinary Digital Publishing Institute: 5405.