USE OF SOCIAL MEDIA FOR ASSESSING SUSTAINABLE URBAN MOBILITY INDICATORS

ALEXANDROS SDOUKOPOULOS, ANASTASIA NIKOLAIDOU, MAGDA PITSIAVA-LATINOPOULOU &
PANAGIOTIS PAPAIOANNOU
Transport Engineering Laboratory, Civil Engineering Department, Aristotle University of Thessaloniki, Greece.

ABSTRACT

Achieving sustainable urban mobility is a complex and multivariate issue that requires constant monitoring and evaluation of the existing situation and possible reconsideration and adjustment of objectives and strategy. The use of indicators is perhaps the most common methodological assessment tool for the sustainable urban mobility level achieved. Key performance indicators can provide in a simple way useful information for complex phenomena in an urban area (i.e. identification of the specific problems and their development over time). Thus, they contribute at a great degree to the decisions made concerning the prioritization of measures and policies toward achieving a goal. However, the use of indicators often constitutes a highly time consuming and costly process due to the large volumes of raw data required for their calculation. In recent years, a solution toward this problem is attempted to be given through the adoption of new technologies and approaches, such as the collection and export of 'big data' from social networks such as Facebook, Twitter, etc. Social networks provide to their users a continuous and enhanced ability for communication, interface and interaction. Such networks are therefore an important potential tool for the promotion of research in the transport sector, as the amount of data generated in their context gives the possibility to analyse and investigate with greater precision critical issues (e.g. trips characteristics) of urban mobility. The present study is an attempt to link the indicators related to sustainable mobility with social networks. The main advantage resulting from the above link, beyond the possibility of a more precise evaluation of the indicators, is to highlight the society's position toward the prioritization of the various transport-related aspects and measures.
Keywords: big data, sentiment analysis, social media, sustainable urban mobility indicators, Twitter

1 INTRODUCTION

Almost half a century ago the United Nations Conference on the Human Environment (1972) [1] had already expressed concern over the followed development models and highlighted the emergence to tackle with the continuous environmental deterioration. Fifteen years later (1987), the World Commission on Environment and Development developed one of the most widely accepted definitions of sustainable development i.e. "meeting the needs of the present without compromising the ability of future generations to meet their own needs" [2], while a few years later (1992) the United Nations, through the Conference on Environment and Development [3] being held in Rio, Brazil promoted further the novel concept of sustainable development. Achieving sustainable development constitutes a complex and challenging issue that requires long-term fundamental changes in a wide range of different aspects including consumption and production patterns, resources management, living standards and mobility patterns. Thus, despite the early introduction of the above-mentioned policy documents over decades ago and the substantial progress being achieved since then, sustainable development still remains an ongoing ambitious goal and the vision for every urban area or country, as many issues have to be further addressed.

Planning toward sustainability constitutes a critical issue particularly in urban areas. Since 2007, the world's urban population exceeds at a great degree the respective rural, while according to the United Nations' estimations, by the year 2030, people living in cities will correspond to the 60% of the world's population and account for the 80% of the global

economic wealth [4, 5]. In this regard, urban areas play a key role in social and economic activities despite the present poor conditions of the environment due to externalities [6]. Since a great share of the externalities derives from the operation of the transport system which comprises a major component of every urban area, it becomes clear that there is a need for adjusting and incorporating the main principles of sustainable development into transport planning [7, 8]. According to the Centre for Sustainable Transportation (1997) and the European Council of Ministers of Transport (2001), sustainable transport planning should aim at a system that "allows the basic access and needs to be met safely and consistently with human and ecosystem health, promotes equity, is affordable and operates efficiently, supports a competitive economy, limits emissions and waste within the planet's ability to absorb them and uses renewable resources at or below their generation rates" [9, 10]. Toward this direction, the constant monitoring and assessment of the current mobility conditions is considered as a necessary process since 'what gets measured gets finally managed', while at the same time an 'issue not clearly defined and measured is also considered as difficult to be improved' [11, 12]. In this framework, indicators due to their numerous advantages have been widely accepted and increasingly implemented during the last years as a methodological tool aiming to capture and evaluate transport sustainability. More specifically, indicators are usually defined as variables or statistics designed to track and assess progress toward achieving either a vision such as sustainable mobility or specific goals and objectives [9, 13, 14]. Indicators identify problems, highlight strengths and weaknesses, contribute to the prioritization of measures and thus facilitate significantly decision-taking and policy-making processes [15, 16]. Moreover, indicators communicate complex phenomena in a simple way being easily understandable by stakeholders, experts and public [17–19]. Indicators are considered as capable of illustrating not only the significant trends but also even the slight changes over time, while they also enable comparisons between different urban areas and jurisdictions [9, 20, 21].

Regarding the indicators' selection process, several authors propose specific criteria for the choice of the appropriate indicators for the development of a relevant system or index. To begin with, indicators should be relevant to policies, objectives and goals that are expected to measure, while they should be also comprehensive by providing useful information concerning the performance in terms of social equity, economic efficiency and environmental integrity [11, 15, 22]. Indicators should be both accessible and available to the public, policy makers and other stakeholders, whereas their structure should be transparent, simple, yet explicitly defined [11, 15, 23–25]. As far as indicators' original data is concerned, the gathered data required for the indicators' calculation must be of high quality and accuracy. At the same time, the data collection process must be affordable and cost-effective i.e. able to be completed within available time, budget and resources [23].

The latter often comprises the weakest point during the indicator selection process, as a result of the large volumes of raw data that several indicators require as well as the local authorities' limited abilities in data gathering due to the lack of efficient and reliable data collection mechanisms [26]. Hence, within the framework of the current research, an alternative methodological approach involving big data deriving from social media such as Twitter is described aiming at estimating sustainable urban mobility indicators that would normally require highly time consuming and costly mobility surveys.

2 THE REVOLUTION OF SOCIAL MEDIA

The rapid development of technology during the last years allowed the evolution of World Wide Web from static websites to the Web 2.0. The basic characteristic that differentiates Web 2.0 platforms and applications, also known as social media, from the classic websites and

blogs is that content is no longer created and published by individuals, but instead is continuously modified by all users. Kaplan and Haenlein (2010) defined social media as "a group of Internet-based applications that are built on the ideological and technological foundations of Web 2.0 and allow the creation and exchange of user-generated content" [27]. Social media facilitate users to interact with each other as well as to share experiences, opinions, knowledge and their locations. Sterne (2010) proposed the classification of different social media into the following seven broad categories according to their scope and content [28]:

- Forums and message boards: exchanging of thoughts, opinions and experiences.
- Review and opinion sites: sharing reviews and opinions, responding to customers' feedback (e.g. TripAdvisor, Yelp).
- Social networks: connecting people and sharing common interests (e.g. Facebook, LinkedIn).
- Blogging: posting news and relative informative material, sharing users' opinions and interests (e.g. LiveJournal, WordPress).
- Microblogging: same as blogging sites yet supporting only short posts (e.g. Twitter).
- Bookmarking: saving, organizing and sharing internet bookmarks with others users (e.g. Del.ici.ous, StumbleUpon).
- Media sharing: uploading and sharing photos and videos (e.g. Flickr, YouTube).

The continuous stream of user-generated content as well as the combination of temporal, spatial and textual data that many of the above-mentioned social media provide, contributed to the creation of a new research field. Researchers belonging to various scientific domains gather and analyze big data deriving from social media in an attempt to come up with interesting patterns and significant findings on a wide range of issues such as human behaviour, public opinions and needs on different issues, schemes and projects as well as in the framework of marketing and political campaigns, etc.

As far as data gathering is concerned, social media mining process is implemented i.e. 'an interdisciplinary process that encompasses techniques from computer science, data mining, machine learning, social network analysis, network science, sociology, ethnography, statistics, optimization, and mathematics' [29].

In this context, mining and analyzing data deriving from social media can contribute to transport planning as well. More specifically, social media-derived data can be used for improving the level of service of a transport mode, for enhancing an existing mobility service or establishing a new one as well as for solving complex problems based on real-time incident reports [30]. Furthermore, the differences between the expected and the perceived results of a project or a service can be assessed [31, 32]. Finally, the abovementioned process can facilitate the evaluation of sustainable urban mobility indicators as described below.

3 METHODOLOGICAL APPROACH

The proposed methodological approach aiming to facilitate the assessment of sustainable mobility indicators with the use of social media consists of three main steps explicitly described in the following sub-sections.

1. Selection of sustainable urban mobility indicators.
2. Selection of specific social media application as a data source.
3. Data mining and analysis.

3.1 Selection of sustainable urban mobility indicators

A list of sustainable urban mobility indicators that their estimation could be potentially based on data deriving from social media was developed, taking into account the main findings of an extended literature review on relevant indicator systems as well as the nature, the specific characteristics and the limitations of the required data. The selected 16 indicators are presented in Table 1. As indicated by the table below, eight of the selected indicators refer to users' degree of satisfaction concerning the infrastructure and operation of alternative to private vehicle modes of transport (i.e. public transport, walking and cycling), while the other eight indicators conceptualize the users' perception toward critical aspects of transport system, such as safety, reliability, affordability etc. It should be pointed out here that the choice of these 16 indicators which constitute a basic sustainable urban mobility indicator system, was made taking into consideration the possibility of using social media as a source for extracting the relevant data.

3.2 Selection of specific social media application as a data source

Regarding the second methodological step, Twitter was selected as the most appropriate social media application for transport-related data mining, based on the following characteristics:

- Twitter is the 14th most visited website on a worldwide basis for year 2016 [33].
- People share thoughts and opinions regarding different issues through Twitter.

Table 1: Sustainable urban mobility indicators that their evaluation could be potentially based on data deriving from social media.

Scope	Indicators
Promotion of public transport (PT)	Level of satisfaction concerning train
	Level of satisfaction concerning metro
	Level of satisfaction concerning tram
	Level of satisfaction concerning bus
	Level of satisfaction concerning taxi
	Level of satisfaction concerning uber
	Perception of PT reliability
	Perception of PT cleanliness
Promotion of non-motorized modes of transport	Level of satisfaction concerning walking conditions
	Level of satisfaction concerning bicycling conditions
Accessible urban environment	Perception of accessibility
Affordable transport system	Perception of transport affordability
Clean urban environment	Perception of emissions and air quality
Safe and secure urban environment	Perception of safety
	Perception of security
Non-congested urban environment	Perception of congestion

- Twitter corresponds to microblogging sites making it suitable for brief exchange of opinions and ideas.
- The collection and classification of data related to a specific topic or subject is considered as easier relative to other social media since hashtags are used.
- Twitter API is considered as more open and accessible in contrast with the API's of other social media e.g. Facebook, since it allows programmers to access large amount of data.
- Opinion mining or sentiment analysis is conducted easily in the context of Twitter, due to the restricted length of posts.
- User-friendly social networking tools like NodeXL extract data from Twitter in a simple and automated way reducing significantly the required time.

3.3 Data mining and analysis

Data mining from social media can be achieved through different methods and pieces of software e.g. Weika, R package and NodeXL. In the framework of the present study, NodeXL was selected since it is considered as an user-friendly, open source software that provides easy access to social media network data streams, calculates advanced network metrics, conducts text and sentiment analysis while also visualizes graph networks [34].

Subsequently, as a pilot study for the implementation of the proposed methodological approach was selected the city of London, which meets the following basic requirements: i) the existence of a large number of transport-related tweets and ii) the great majority of tweets is being written in English, a fact that facilitates further the analysis, as NodeXL does not perform by default specific text metrics in other languages. It should be pointed out here that the main scope of this pilot implementation comprises the investigation of the strengths as well as the weaknesses and limitations of the proposed methodological approach rather than the provision of results in absolute terms.

Since the selection of appropriate keywords constitutes one of the most critical components of data mining process, several attempts have been made before ending up in a set of keywords being associated with the selected sustainable urban mobility indicators. It should be also pointed out here that transport-related keywords yet of low frequency of use were excluded from the selected set, while data mining process was conducted twice in order to collect tweets that either include the keyword inside the text or as a hashtag. The above-mentioned approach resulted to a significant number of collected tweets accounting for 18.000 per keyword. However, since the percentage of automated or manually geo-located tweets proved to be extremely low ranging from 0–2%, the collection of a satisfactory number of relevant tweets referring to London could not be achieved. In order to cope with this inconsistency, new keywords including the spatial information either inside the tweet or as a hashtag were sought. Hence, the final data mining was conducted during the period from 01/03/2017 to 15/03/2017 using only one keyword namely 'London transport'. The use of the above-mentioned keyword, despite being only one, led to the collection of all data required to evaluate the 16 selected sustainable urban mobility indicators while simultaneously ensured that the spatial reference refers to London. The total number of collected tweets amounted to 17,233 deriving from 11,926 unique Twitter users. The overall metrics are presented in Table 2. According to the table below, the number of tweets per user is rather small indicating lack of many predominant users and therefore diversity in the collected tweets. As far as geolocation is concerned, the percentage of geo-located tweets remains extremely low and thus prevents any further grouping of tweets based on their exact location inside London.

Table 2: Overall metrics of finally collected tweets referring to the keyword 'London transport'.

Basic metrics	Results
Total number of tweets	17,233
Total number of users	11,926
Number of tweets per user	1.44
Percentage of geolocated tweets	0.23%
Percentage of tweets with publication of location	2.11%
Total number of languages in which tweets were published	18
Percentage of tweets written in English	94.98%
Percentage of tweets written in French	0.66%
Percentage of tweets written in Spanish	0.11%

The collected tweets were grouped into 16 categories referring to the 16 selected sustainable urban mobility indicators (presented in Table 1). Accordingly, sentiment analysis was conducted in order to evaluate the sustainable mobility indicators. Sentiment analysis is considered as a useful tool for investigating and understanding public's opinions regarding different subjects [35]. It can be described as an 'automated process for extracting opinions, sentiment and subjectivity from text regarding an entity or an event using natural language processing, statistical and machine learning techniques' [36]. The objective of sentiment analysis is the classification of text into a number of predefined sentiment clusters. One of the most widely accepted methods for sentiment analysis consists of Sentiment Lexicons, a method corresponding to the 'Lexicon-based' category. The lexicon includes words marked either as negative or positive depending on the sentiment that they express. This method is used to analyze emotions in text and thus assign a specific sentiment score to the text [37].

In the context of current study, sentiment analysis was conducted using NodeXL. More specifically, frequency as well as the percentage of sentiment words included in the examined tweets were measured. It should be noted here that tweets were treated as a set of unrelated words without considering the order and the grammatical properties; in other words, tweets were handled as a 'bag of words'. Calculation of sentiment was based as described before on predefined sentiment lexicons containing words that express different sentiments. The predefined categories were defined as positive, negative and neutral [38, 39]. Based on the above-mentioned, each tweet was assigned to the predominant sentiment category.

4 PRESENTATION OF THE MAIN RESULTS

The results of the pilot implementation, although they derive from a not comprehensive assessment of the performance of London's transport system, highlighted some interesting findings. However, comparing the findings of the selected sustainable urban mobility indicators with the results of relevant measurements e.g. the perceived level of congestion versus the on-field measurement of congestion, exceeds the scope of the current study.

As far as transport-related terms are concerned, Londoners users of Twitter seem to mention more frequently terms such as 'tube', 'uber', 'car' and 'train' rather than 'bus' 'walking' and 'bicycle'. Accordingly, the most common terms included in the collected tweets were visualized through IBM Wordcloud software and are presented in Fig. 1 below.
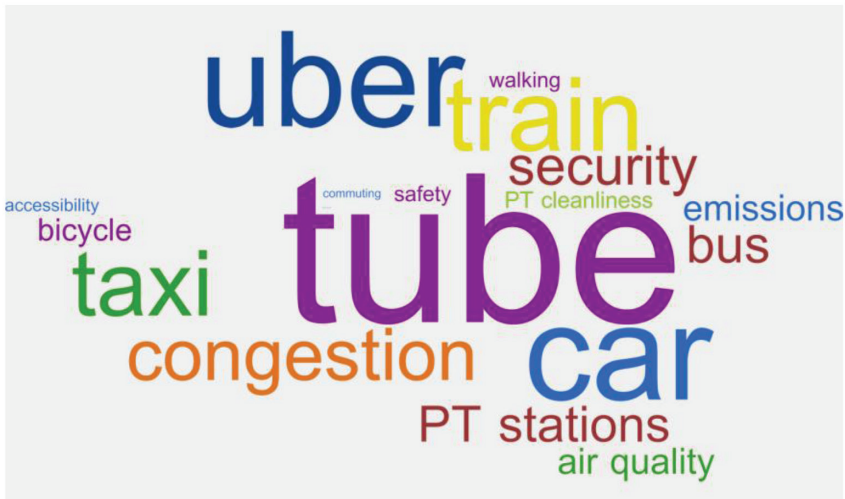
Figure 1: Visualization of the most common keywords included in the collected tweets through IBM Wordcloud software.

A sentiment analysis was also conducted concerning the evaluation of sustainable urban mobility indicators, where it was found that indicators referring to the perceived level of accessibility and affordability as well as the level of bicycle's conditions satisfaction correspond to the most positive sentiment or evaluation. On the other hand, the level of satisfaction concerning the operation of tram, the perceived PT reliability and the perceived level of congestion are considered negatively.

The visualisation of sentiment analysis as well as the results of the evaluation of the sustainable urban mobility indicators are presented in Figs. 2 and 3, respectively. More specifically, Fig. 2 presents the generic sentiment calculated for each indicator through a contemporary
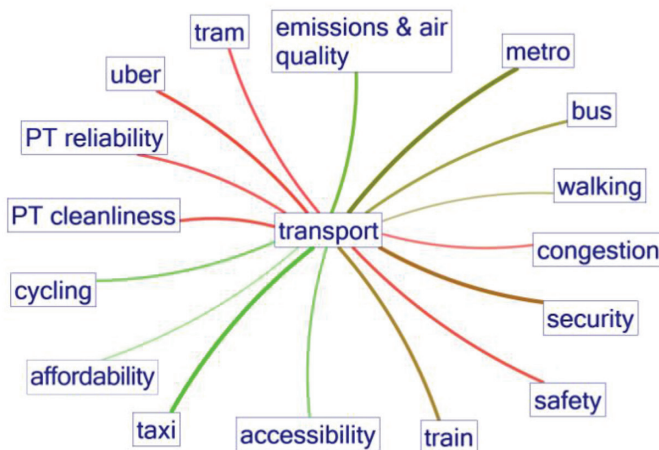

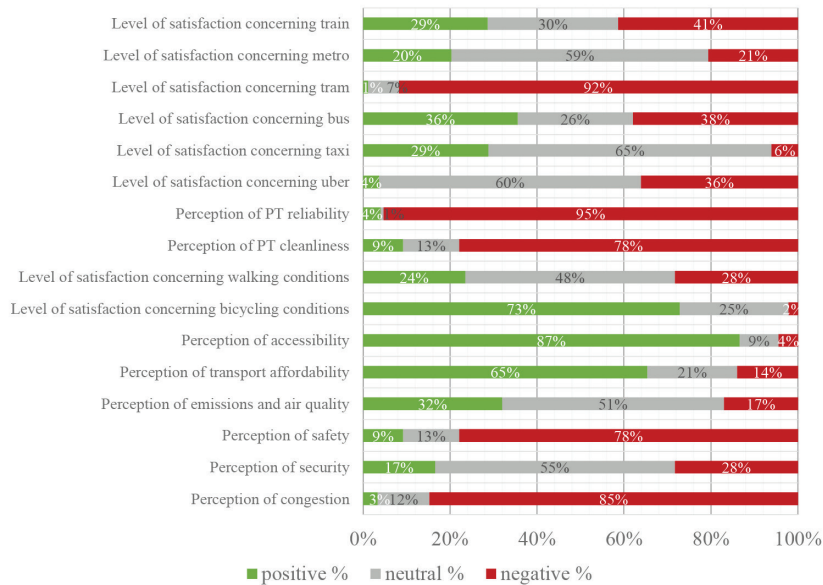Figure 2: Visualization of the sentiment analysis results.

Figure 3: Evaluation of the selected sustainable urban mobility indicators through the results of the conducted sentiment analysis.

presentation method. The colors vary from green for the positive sentiment to red for the negative sentiment and the thickness as well as the visibility of the edges correspond to the total number of tweets related to each indicator. Regarding Fig. 3, the percentages referring to the positive, negative and neutral category are presented explicitly for each indicator.

As indicated by the figures above, the overall predominant sentiment is considered to be negative. This could be explained by the fact that social media users tend to post more often for commenting negative issues or aspects. Moreover, Lexicon-based methods such as the one used in the present study, by default tend to recognize more easily and categorize accordingly the extreme negative tweets due to the intense way of expressing when a problematic issue is reported. Furthermore, some of the selected sustainable urban mobility indicators are related with terms often used in a negative manner such as congestion, safety etc. resulting to even more negative findings.

5 CONCLUDING REMARKS

Indicators constitute a significant methodological tool for assessing current mobility conditions and promoting the achievement of sustainable urban mobility. However, the large volume of raw data often required for evaluating a large number of sustainable urban mobility indicators highlight the emergence for implementing contemporary, low cost and time-efficient data collection methods. In this context, an alternative methodological approach was developed involving big data deriving from social media such as Twitter. The scope of the above-mentioned approach is the estimation of sustainable urban mobility indicators referring to the perceived level of satisfaction regarding different modes or transport-related aspects.

The proposed methodological approach was implemented as a pilot in London in order to highlight its strengths as well as its potential weaknesses and limitations. More specifically,

the vast amount of data being available on social media can be extracted automatically and at a limited budget, especially using software like NodeXL that do not require programming skills. The data deriving from social media correspond to a considerable share of population. This share although does not include specific population classes such as the elderly, is expected to become more and more representative in the years ahead. As far as data deriving from Twitter is concerned, a weakness that was observed is that tweets often lack spatial information since users rarely activate the geolocation feature. Hence, the selection of relevant keywords containing spatial reference is proposed. Regarding the assessment of the selected sustainable urban mobility indicators, sentiment analysis was conducted. Sentiment analysis classifies tweets into a number of predefined sentiment clusters illustrating thus the public opinion toward an issue. However, sentiment analysis is considered as challenging since tweets often include irony or sarcasm pointing out misleading findings. However, despite all the above-mentioned limitations it could be argued that the current study constitutes a first step toward the use of social media as a data source for the assessment of sustainable urban mobility indicators. Further research on the respective issue is necessary in order to highlight, beyond the possibility of a more precise and efficient evaluation of indicators, the society's position toward the prioritization of the various transport-related aspects and measures.

ACKNOWLEDGEMENTS

REFERENCES

[1] United Nations, *Report of the Conference on the Human Environment*. United Nations Publication, 1972.

[2] United Nations, *Report of the World Commission on Environment and Development*. United Nations Publication, 1987.

[3] United Nations, *Report of the World Commission on Environment and Development, Agenda 21: Programme of Action for Sustainable Development - Rio Declaration*. United Nations Publication, 1992.

[4] United Nations-Habitat, *The state of the world's cities report 2006/2007: 30 Years of Shaping the Habitat Agenda*, Earthscan: London, 2006.

[5] Muneer, T., Celik, A.N. & Caliskan, N., Sustainable transport solution for a medium-sized town in Turkey - A case study. *Sustainable Cities and Society*, **1**(1), pp. 29–37, 2011.
http://dx.doi.org/10.1016/j.scs.2010.08.004

[6] Mori, K. & Christodoulou, A., Review of sustainability indices and indicators: Towards a new City Sustainability Index (CSI). *Environmental Impact Assessment Review*, **32**(1), pp. 94–106, 2012.
http://dx.doi.org/10.1016/j.eiar.2011.06.001

[7] Litman, T.A., Sustainable Transportation Indicators - A Recommended Research Program for Developing Sustainable Transportation Indicators and Data. *Transportation Research Board Annual Meeting*, 2009.

[8] Tafidis, P., Sdoukopoulos, A. & Pitsiava-Latinopoulou, M., Sustainable urban mobility indicators: policy versus practice in the case of Greek cities. *Proceedings of the 3rd Conference on Sustainable Urban Mobility*, pp. 340–347, 2016.

[9] Gilbert, R., Irwin, N., Hollingworth, B. & Blais, P., Sustainable Transportation Performance Indicators. *Transportation Research Board Annual Meeting*, 2003.

[10] European Union, *Minutes of the 2340th meeting of the Council of Ministers*, Luxembourg, 2001.

[11] Zheng, J., Garrick, N.W., Atkinson-Palombo, C., McCahill, C. & Marshall, W., Guidelines on developing performance metrics for evaluating transportation sustainability. *Research in Transportation Business & Management*, **7**, pp. 4–13, 2013.
http://dx.doi.org/10.1016/j.rtbm.2013.02.001

[12] Böhringer, C. & Jochem, P.E.P., Measuring the immeasurable - A survey of sustainability indices. *Ecological Economics*, **63**(1), pp. 1–8, 2007.
http://dx.doi.org/10.1016/j.ecolecon.2007.03.008

[13] Litman, T.A., *Well Measured - Developing Indicators for Sustainable and Livable Transport Planning*. Victoria Transport Policy Institute, 2015.

[14] Henderson, H., *Building A Win-Win-World*, Berrett-Koehler: San Francisco, 1996.

[15] Dobranskyte-Niskota, A., Perujo, A. & Pregl, M., *Indicators to assess sustainability of transport activities Part 1: Review of the Existing Transport Sustainability Indicators Initiatives and Development of an Indicator Set to Assess Transport Sustainability Performance*. European Commission, Joint Research Centre, 2007.

[16] Costa, M.S., Silva, A.N.R. & Ramos, R.A.R., Sustainable urban mobility: a comparative study and the basis for a management system in Brazil and Portugal. *WIT Transactions on The Built Environment*, vol. 77, WIT Press: Southampton and Boston, pp. 323–332, 2005.

[17] European Environment Agency, *EEA core set of indicators: Guide*. EEA Technical Report, Copenhagen, 2005.

[18] Eckersley, R., *Perspectives on progress: Is life getting better? Resource Futures Program, Working Paper Series 97/27*. Resource Futures Program, Commonwealth Scientific and Industrial Research Organisation, Canberra, 1997.

[19] Organization of Economic Coordination and Development, *Environmental Indicators. OECD Core Set of Indicators for Environmental Performance Reviews.* Environment Monographs no 83, OECD: Paris, 1993.

[20] Joumard, R., Gudmundsson, H. & Folkeson, L., Framework for assessing indicators of environmental impacts in the transport sector. *Transportation Research Record: Journal of the Transportation Research Board,* **2242**, pp. 55–63, 2011.
http://dx.doi.org/10.3141/2242-07

[21] Tanguay, G.A., Rajaonson, J., Lefebvre, J.F. & Lanoie, P., Measuring the sustainability of cities: An analysis of the use of local indicators. *Ecological Indicators*, **10**(2), pp. 407–418, 2010.
http://dx.doi.org/10.1016/j.ecolind.2009.07.013

[22] Litman, T.A., *Sustainable transportation indicator data quality and availability*. Victoria Transport Policy Institute, 2009.

[23] Joumard, R. & Gudmundsson, H., (eds), *Indicators of environmental sustainability in transport: An interdisciplinary approach to methods*, INRETS Report: Bron France, 2010.

[24] Santos, A.S. & Ribeiro, S.K., The use of sustainability indicators in urban passenger transport during the decision-making process: the case of Rio de Janeiro, Brazil. *Current Opinion in Environmental Sustainability*, **5**(2), pp. 251–260, 2013.
http://dx.doi.org/10.1016/j.cosust.2013.04.010

[25] Miller, H.J., Witlox, F. & Tribby, C.P., Developing context-sensitive livability indicators for transportation planning: a measurement framework. *Journal of Transport Geography*, **26**, pp. 51–64, 2013.
http://dx.doi.org/10.1016/j.jtrangeo.2012.08.007

[26] Burggraf, K., *Why is Monitoring and Evaluation a Challenge in Sustainable Urban Mobility Planning?*. Report of CH4LLENGE programme, 2014.

[27] Kaplan, A.M. & Haenlein, M., Users of the World, Unite! The challenges and opportunities of social media. *Business Horizons*, **53**(1), pp. 59–68, 2010.
http://dx.doi.org/10.1016/j.bushor.2009.09.003

[28] Sterne, J., *Social Media Metrics: How to Measure and Optimize Your Marketing Investment*, John Wiley & Sons: Chichester UK, 2010.

[29] Zafarani, R., Abbasi, M.A. & Liu, H., *Social Media Mining: An introduction*, Cambridge University Press: Cambridge UK, 2014.

[30] Kenyon, S. & Lyons, G., The value of integrated multimodal traveller information and its potential contribution to modal change. *Transport Research Part F: Traffic Psychology Behavior*, **6**(1), pp. 1–21, 2003.
http://dx.doi.org/10.1016/s1369-8478(02)00035-9

[31] Eboli, L. & Mazzulla, G., A methodology for evaluating transit service quality based on subjective and objective measures from the passenger's point of view. *Transport Policy*, **18**(1), pp. 172–181, 2011.
http://dx.doi.org/10.1016/j.tranpol.2010.07.007

[32] Mazzulla, G. & Forciniti, C., Spatial association techniques for analysing trip distribution in an urban area. *European Transport Research Review*, **4**(4), pp. 217–233, 2012.
http://dx.doi.org/10.1007/s12544-012-0082-9

[33] Alexa Internet Inc., *The top 500 sites on the web*. www.alexa.com/topsites, last access February 20, 2017.

[34] Hansen, D.L., Shneiderman, B. & Smith, M.A., *Analyzing Social Media Networks with NodeXL: Insights from a Connected World*, Morgan Kaufmann: Burlington USA, 2011.

[35] Pang, B. & Lee, L., Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, **2**(1–2), pp. 1–135, 2008.
http://dx.doi.org/10.1561/1500000011

[36] Park, A. & Paroubek, P., Twitter as a corpus for sentiment analysis and opinion mining. *Proceedings of International Conference on Language Resources and Evaluation*, pp. 17–23, 2010.

[37] Sood, S., Owsley, S., Hammond, K. & Birnbaum, L. *Reasoning through search: a novel approach to sentiment classification*. Technical Report, 2007.

[38] Hu, M. & Liu, B., Mining and summarizing customer reviews. *Proceedings of the 10th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 168–177, 2004.

[39] Liu, B., Hu, M. & Cheng, J., Opinion Observer: Analyzing and Comparing Opinions on the Web. *Proceedings of the International World Wide Web Conference Committee (IW3C2)*, 2005.