

86 Critical comparison of social media and other user-generated geographic information as a source of visitor information – lessons learned in the SoMeCon-project

Vuokko Heikinheimo^{1,2}, Tuuli Toivonen², ¹Finnish Environment Institute SYKE, Finland. ²Digital Geography Lab, University of Helsinki, Finland

Up-to-date information about outdoor recreation experiences is important for the planning and management of national parks and other outdoor destinations. User-generated data such as geotagged social media posts and GPS-tracks shared via sports applications have emerged as potential new data sources to complement on-site counters and surveys. There are considerable amounts of geographic information available from digital platforms and mobile devices representing the movements, activities and preferences of visitors, and these data have been increasingly used for studying visits to national parks and green spaces..

This presentation draws together our findings from the Social Media Data for Conservation Science –project SoMeCon (2016-2021). Our main objectives were to 1) gain methodological understanding about social media and other user-generated data sets as a source of geographic information, and 2) to provide new information about the spatial and temporal patterns of human activities in national parks and green spaces. We compared social media data to official visitor statistics from Finnish and South African national parks, and social media to other sources of user-generated geographic information (sports app data, mobile network data, PPGIS data) from urban green spaces in Helsinki, Finland.

We provide an overview of methods that are useful for analyzing the spatial and temporal dimensions of user-generated data as well as the rich visual and textual content (Toivonen et al. 2019). We also discuss the limitations related to data quality, limited access as well as privacy issues. On the one hand, anonymized and aggregated data products are one way forward regarding ethical analysis of user-generated geographic information. On the other hand, data minimization and pseudonymization already at the data collection phase helps to overcome some issues related to protecting personal

information (Di Minin et al. 2021; Sandbrook et al. 2021).

Comparisons between social media data and official visitor information support the use of social media data as an indicator of park visits and visitor preferences to some extent. In popular parks, temporal changes in social media posting activity reflect changes in measured visits, while in less visited and remote parks the social media posting patterns are more sporadic and irregular (Tenkanen et al. 2017). We found similar trends in both social media content and visitor surveys regarding popular activities and visited places (Heikinheimo et al. 2017; Hausmann et al. 2018). Social media photos may reveal also seasonal patterns in the landscape (like snow cover) and how the activities change by the season. Automated analysis approaches such as computer vision methods offer new opportunities for summarizing such information from vast quantities of data (Väisänen et al. 2021).

Our comparisons among social media data, sports application data, mobile phone data and participatory geographic information from urban green spaces highlights the questions to which each data source can contribute to (Heikinheimo et al. 2020). Data from social media platforms reflects enjoying and being in nature, sports app data and mobile network data reflect more precise movements during a visit or a daily commute. PPGIS data are useful for capturing values and preferences more in depth from a targeted sample. Each data source has its limitations regarding data access, spatial and temporal extent and metadata availability.

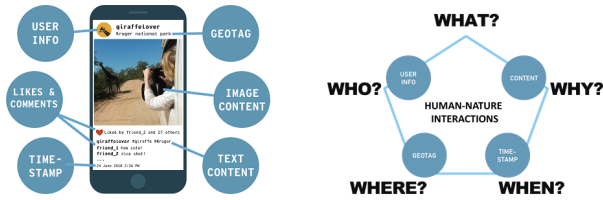


Figure 1. Various elements of user-generated geographic information offer novel perspectives for understanding the spatio-temporal patterns of human activities in nature.

Overall, user-generated geographic information offers valuable insights about where, when and how people use and value nature (Figure 1). Combining information from several sources provides a more

comprehensive understanding of green space use and preferences. In areas outside official visitor monitoring schemes, user-generated data might be the best available information about the movement and activities of people. Even if these user-generated data sources are imperfect measures of nature's contributions to people, they are still better than ignoring the value of to human well-being completely. Novel data sources such as social media photographs or traces from mobile network data should be used to complement, not to replace, existing high-quality visitor monitoring schemes.

References

Di Minin et al. 2021. <https://doi.org/10.1111/cobi.13708>. Hausmann et al. 2018. <https://doi.org/10.1111/conl.12343>. Heikinheimo et al. 2020. <https://doi.org/10.1016/j.landurbplan.2020.1038>. Heikinheimo et al. 2017. <https://doi.org/10.3390/ijgi6030085>. Sandbrook et al. 2021. <https://doi.org/10.1111/csp2.374>. Tenkanen et al. 2017. <https://doi.org/10.1038/s41598-017-18007-4>. Toivonen et al. 2019. <https://doi.org/10.1016/J.BIOCON.2019.01.023>. Väisänen et al. 2021. <https://doi.org/10.1111/cobi.13704>