AUTHOR Andrea Soriano-Redondo, Ricardo A. Correia, Enrico Di Minin

TITLE Social media data can inform the Global Biodiversity Framework

YEAR 2023

DOI https://doi.org/10.1016/j.biocon.2023.109993

VERSION Author's accepted manuscript

CITATION Andrea Soriano-Redondo, Ricardo A. Correia, Enrico Di Minin, Social media data can inform the Global Biodiversity Framework, Biological Conservation, Volume 281, 2023, 109993, ISSN 0006-3207, https://doi.org/10.1016/j.biocon.2023.109993. (https://www.sciencedirect.com/science/article/pii/S0006320723000939)
Keywords: Global Biodiversity Framework; Biodiversity monitoring; Social media data

Social media data can inform the Global Biodiversity Framework

Andrea Soriano-Redondo[1,a*], Ricardo A. Correia[1,2,3,a*], Enrico Di Minin[1,3,4,a*]

1. Helsinki Lab of Interdisciplinary Conservation Science (HELICS), Department of Geosciences and Geography, University of Helsinki, 00014, Helsinki, Finland.

2. Biodiversity Unit, University of Turku, 20014 Turku, Finland.

3. Helsinki Institute of Sustainability Science (HELSUS), University of Helsinki, 00014 Helsinki, Finland.

4. School of Life Sciences, University of KwaZulu-Natal, 4041 Durban, South Africa.

[a] All authors contributed equally.

*Correspondence: Andrea Soriano-Redondo andrea.sorianoredondo@helsinki.fi +358294150770; Ricardo A. Correia ricardo.correia@helsinki.fi; Enrico Di Minin enrico.di.minin@helsinki.fi

Acknowledgements

A.S.R. was supported by a Marie Skłodowska-Curie Actions Postdoctoral Fellowship (grant agreement 101022521). R.A.C. acknowledges personal funding from the Academy of Finland (grant 348352) and from the KONE Foundation (grant 202101976). E.D.M. thanks the European Research Council (ERC) for funding under the European Union's Horizon 2020 research and innovation programme (grant agreement 802933).

In their Perspective, Orr et al. (2022) state the need for quality and up-to-date biodiversity data and outline a vision for automation and synthesis of conservation data resources to inform decision-making. The authors emphasize existing biases and limitations in currently available species distribution data, including taxonomic, spatial, and temporal gaps. While we agree with the authors on the need for better coverage, more reliable data, and synthesis efforts to effectively inform conservation decision-making at multiple scales, we believe they have neglected an important source of information that can be used for monitoring the status and distribution of biodiversity, social media data.

We believe social media data can yield novel and (near) real-time information on the distribution of biodiversity, and even the pressures affecting it. Data collected from social media platforms provide unique insights on the current distribution of species (Toivonen et al., 2019), can be used to explore species range shifts in response to global changes and have even led to the discovery of new species. Data collected from social media can also help fill other biodiversity information gaps related to species traits and interactions of phenological dynamics (Jarić et al., 2020). Furthermore, social media data can be used as an additional and independent source of information for recursive cross-checking across data types, especially to provide novel locations for species for which only old data are available. If collected following existing data privacy and protection safeguards, these data can potentially be used without exposing social media users and in full respect to privacy regulations (Di Minin et al., 2020). Social media data can be particularly useful in areas where biodiversity information is currently lacking or in socio-ecological landscapes that support people and nature, and it can act as an important complement to the much-needed digitization efforts of historical biodiversity records suggested by the authors.

While we see great potential in using social media data to enhance biodiversity knowledge, we also recognize that including such data in biodiversity monitoring efforts faces challenges. For example, similar to efforts developed for citizen-science data, social media data would require specific protocols for expert validation. Data validation protocols could be implemented at

distinct stages. One option is that data are filtered and validated by experts before they are made available for use in biodiversity repositories, such as the Global Biodiversity Information Facility (GBIF). Another option is to submit the data without prior validation but including quality flags alerting on the need for further refinement by experts. Finally, ongoing developments in machine learning can facilitate a mixed approach, providing a first level of data filtering and validation to ensure minimum quality standards in shared data until it can be consulted and validated by experts. Although social media data can suffer from accessibility restrictions and multiple biases, including limited geographic and temporal coverage, tapping into multiple data sources and the development of new modelling approaches that consider uncertainty from data biases can help address these issues (Tessarolo et al., 2021). Efforts to increase individual participation in data generation, and data and resource sharing by social media companies would also be important to minimize these limitations.

Ultimately, we call for social media data to be considered and included into efforts to develop a unified monitoring framework that integrates multiple data sources to measure progress towards the Kunming-Montreal Global Biodiversity Framework and the United Nation Sustainable Development Goals. Social media data can help fill information gaps in countries where resources for monitoring and citizen science enthusiasts are less readily available. Lack of funding for conservation will continue to be a limiting factor, so cost-effective means of collecting information to monitor progress towards meeting global to local biodiversity targets are needed until more systematic monitoring efforts can be put in place. The integration of social media data within a broader framework for data collection and synthesis from multiple sources can help address this issue and provide novel insights for multiple stakeholders.

References

Di Minin, E., Fink, C., Hausmann, A., Kremer, J., Kulkarni, R., 2021. How to address data privacy concerns when using social media data in conservation science. Conservation Biology 35, 437–446. https://doi.org/10.1111/cobi.13708

Jarić, I., Correia, R.A., Brook, B.W., Buettel, J.C., Courchamp, F., Di Minin, E., Firth, J.A., Gaston, K.J., Jepson, P., Kalinkat, G., Ladle, R., Soriano-Redondo, A., Souza, A.T., Roll, U., 2020. iEcology: Harnessing Large Online Resources to Generate Ecological Insights. Trends in Ecology & Evolution 35, 630–639. https://doi.org/10.1016/j.tree.2020.03.003

Orr, M.C., Hughes, A.C., Costello, M.J., Qiao, H., 2022. Biodiversity data synthesis is critical for realizing a functional post-2020 framework. Biological Conservation 274, 109735. https://doi.org/10.1016/j.biocon.2022.109735

Tessarolo, G., Ladle, R.J., Lobo, J.M., Rangel, T.F., Hortal, J., 2021. Using maps of biogeographical ignorance to reveal the uncertainty in distributional data hidden in species distribution models. Ecography 44, 1743–1755. https://doi.org/10.1111/ecog.05793

Toivonen, T., Heikinheimo, V., Fink, C., Hausmann, A., Hiippala, T., Järv, O., Tenkanen, H., Di Minin, E., 2019. Social media data for conservation science: A methodological overview. Biological Conservation 233, 298–315. https://doi.org/10.1016/j.biocon.2019.01.023