From Tweets to Tings: Dissimilarity in Recorded Species Community Between Social Media Platforms and Implications for Resource-Limited Conservation

Adam G. Hart*[1], Elliot Tripp[1] and Anne E. Goodenough[1]

1. School of Education and Applied Sciences, Francis Close Hall, University of Gloucestershire, Cheltenham, GL50 4AZ

   *Corresponding Author: ahart@glos.ac.uk ORCID 0000-0002-4795-9986

Abstract

Effective wildlife conservation relies on surveying species, yet traditional methods are often financially- and labour-intensive. Leveraging social media posts is potentially a more efficient alternative. Here, community assemblages derived from mammal sightings in Kruger National Park, South Africa, over the same four-month period were compared for five social media platforms: Instagram, X (formerly Twitter), Facebook, Flickr, and Kruger Latest Sightings (a bespoke platform). In total, 7,860 unique sightings of 45 species were recorded. Latest Sightings was the most comprehensive and consistent platform (37 species, 82%), while X provided the least complete species assemblage (28 species, 62%). There were significant inter-platform differences in species community composition but records of charismatic megafauna including lion and leopard were extremely high across platforms (37% of total sightings), highlighting user biases. Use of multi-platform data proved superior to any single platform in isolation, underscoring the need for a diversified approach. Our findings demonstrate the feasibility and challenges of using social media for wildlife monitoring, emphasizing the importance of understanding platform-specific user behaviours. This study contributes to the growing body of evidence on iEcology and the use of social media in conservation efforts, particularly in resource-limited contexts, and provides advice for integration into ecological monitoring strategies.

Keywords

iEcology, social media, conservation, ecological monitoring, surveying

Statements and Declarations

This study was approved by the Ethics Committee of the University of Gloucestershire (Ref: NSSS.EV1123.01).

The authors declare no competing financial or non-financial interests, or any other conflicts of interest relevant to this study.

Introduction

Effective wildlife conservation relies on monitoring species communities, including quantifying presence and abundance, as well as population characteristics such as sex and age structure (Goodenough and Hart, 2017). Obtaining monitoring data can involve considerable financial cost and be labour intensive. Consequently, despite its importance, ecological monitoring is often resource-limited and is undertaken less frequently, less intensively, or for a shorter period than would be adequate (Field et al., 2005; Grantham et al., 2009).

To overcome resource limitations, ecologists are turning to technological approaches (such as camera traps and bioacoustic recorders, now often paired with machine learning to enable automated species recognition) and unpaid volunteer recorders (through citizen science or use of ad-hoc records) to undertake species monitoring (McKinley et al., 2017; Sumner et al., 2019; Chhaya et al, 2021; Hart et al., 2022; Mitterwallner al., 2023). Such approaches can be effective and robust, but still require resources, including physical assets, data recorders, development of bespoke apps, hosting fees, and staff costs for setup, data management, and analysis.

A less resource-intensive technological approach is to take advantage of indirect information, such as social media posts. Wildlife sightings are often posted on social media platforms and data useful for establishing species community composition and spatiotemporal change can sometimes be harvested (di Minin et al., 2015; Hart et al., 2018; Ghermandi and Sinclair, 2019; Toivonen et al., 2019, Edwards et al., 2021). This has led to the development of a research approach termed iEcology, which involves quantifying "patterns and processes in the natural world using data accumulated in digital sources collected for other purposes" (Jarić et al., 2021; Pernat et al., 2024). This has been used, for example, to map distribution of bees in the UK (Stafford et al., 2010), quantify migratory phenology of fish and flowering plants in North America (Mittermeier, et al., 2019; Breckheimer et al., 2020), enhance biodiversity understanding and inform conservation planning, for example, in Bangladesh (Chowdhury et al., 2023; Chowdhury et al., 2024), track changes in species distribution of Banded Demoiselle (*Calopteryx splendens*) in Britain (O'Neill et al., 2023), track illegal trade in Mexico (Salas-Picazo et al., 2023), and determine raptor diet in South America and Africa (Naude et al., 2019; Berryman and Kirwan, 2021). In particular, "mining" of the microblogging platform X (formerly Twitter), either manually or automatically via Application Programming Interfaces (APIs), can provide valuable incidental data relevant to community ecology, including biodiversity occurrence, phenology, and behaviour (e.g. Daume, 2016; Hart et al., 2018; Edwards et al., 2022).

National Parks in Southern Africa attract considerable numbers of wildlife-oriented visitors, both locals and overseas tourists, hoping to see species such as elephant (*Loxodonta africana*), giraffe (*Giraffa camelopardalis*) and zebra (*Equus zebra*). African megafauna species are often considered iconic and charismatic (di Minin et al., 2013), but many are of conservation concern, including lion (*Panthera leo*; vulnerable), leopard (*P. pardus*; vulnerable) and black rhinoceros (*Diceros bicornis*; critically endangered). Casual observation of social media platforms shows that visitors to National Parks frequently post sightings of species, including descriptive text and photographs, a view reinforced by recent descriptive studies of wildlife encounters based on social media posts in South Africa (Mangachena et al., 2023) Furthermore, some National Parks have dedicated, actively-managed social media apps that allows users to report their latest wildlife sightings. Social media platforms differ in their overall "tone", user base and intended audience, as well as the typical balance between images and text (e.g. Laor, 2022; Masciantonio et al., 2021). Such factors have the potential to influence the species about which users decide to post, and thus the species community being reported.

With African National Parks becoming increasingly underfunded (Lindsey et al., 2021), making use of tourist social media posts as part of an iEcology approach to monitoring is attractive (Haussman et al., 2017). However, to be effective it is necessary to establish baseline information on how different platforms might differ in terms of species reported. In this study, we examine the use of five popular social media platforms for posts of species sighted in Kruger National Park, South Africa (Instagram, X, Facebook, Flickr and the bespoke app Kruger Latest Sightings). We harvest wildlife sightings from posts made during the same four-month period in 2023 for all platforms and examine the differences in species (and thus ecological communities) reported. This appears to be the first direct comparison of multiple social media platforms in terms of the species community composition recorded in any setting. We discuss similarities and differences in the light of the potential for each platform to provide useful, cost-effective passive monitoring data for on-going conservation efforts.

Methods

Social media platforms

Posts on five social media platforms were analysed during this research. (1) Kruger Latest Sightings: a bespoke app developed in 2011 to allow users to share information on wildlife sightings (termed "tings") with other visitors, including species, date, time, location and visibility. (2) Instagram: a generic platform launched in 2010 with an identity built upon users sharing smartphone images of events with "followers". (3) Facebook: launched in 2004 and now the largest social media platform by userbase, with users sharing text- and image-based content with "friends". (4) X (formerly Twitter): generic microblogging platform that launched in 2007 whereby users post publicly-available messages known as "tweets" (typically 280 characters plus an optional image, although paying subscribers can post longer tweets and all users can post linked tweets as a "thread"). (5) Flickr: a photograph sharing platform that debuted in 2004 and focused on users posting high-quality images with secondary text descriptions.

Data collection

Manual searches of the five social media platforms were undertaken to retrieve posts about wildlife sightings in Kruger National Park made over a 4-month (17-week) period between mid-February and mid-June 2023. Posts relating to mammals (excluding bats and rodents where identification was often uncertain) were retrieved using preexisting personal accounts to search platforms.

To ensure that searching social media platforms was systematic, and thus that post retrieval was replicable and consistent throughout the study period, we followed the precepts of the PRISMA framework often used for systematic retrieval of literature or data for use in meta-analysis (Page et al., 2021). On each social media platform, for each week of the study period, the following PRISMA steps were followed:

(1) Identification of candidate posts based on key search terms entered manually: we used eight generic search terms - Kruger, Kruger National Park, KNP, Kruger Wildlife, Kruger Safari, Kruger Sightings, Latest Sightings Kruger, Kruger Latest – but the exact search framework was customised for each social media platform. On Facebook and X/Twitter, which support full text searching as well as searching of hashtags, each search term was entered using standard prose, within quotation marks, and as hashtags. In addition, on Facebook, posts were also collated directly from the timeline of the two largest wildlife sighting groups for Kruger: Kruger Sightings (227,000+ members) and Kruger Wildlife Sightings (46,000+ members). On

Instagram, where the search function is restricted to tags rather than full text, search terms were used in hashtag format only; this platform also supported searching within “Kruger National Park” as a location with the search filter set to “recent” to retrieve posts as a timeline, with the most recent post at the top. On Flickr, which has an advanced search facility, search terms were used within full text (which included tags by default) with the “date taken” range being set to the dates at the start and end of each study week. For the Latest Sightings app, no searching was undertaken as all posts were examined using the in-app timeline function.

(2) Screening of candidate posts: posts from all platforms were assessed for relevancy with non-relevant posts (e.g. those not about wildlife; historical reminiscence) removed.

(3) Extracting data from posts that remained in scope: using information from the post text and any imagery, the following data were manually extracted from each post: platform; sighting date (if not provided, the date of the post was assumed to be the date of the sighting); location; sighting time (where given); and species. Descriptions, locations, and any images were visually compared to try to identify situations where the same sighting had been recorded on multiple social media platforms by the same (or different) users. To reduce pseudoreplication for multi-post sightings, the first post chronologically was retained while other posts (or reposts) of the same sighting were removed. It is, however, recognised that this process reduced, rather than removed, pseudoreplication because, in instances where two posts of the same sighting recorded key information (location, time, number of animals, etc.) differently, there was no sensible way for this to be identified. Ethical clearance was provided by University of Gloucestershire (Ref: NSSS.EV1123.01), which stipulated that no personal information relating to the original poster should be extracted and no aspect of data mining should violate platform-specific end-user agreements.

We recognise that our protocol might have missed some posts that would have been within the scope of this research, however, the process adopted is systematic, standardised and transparent without obvious bias in relation to post type or author demographics.

Data analysis

All data analysis was undertaken in the freeware package PAST (PALeontological STatistics) (Hammer, et al., 2001) using the distance metric Bray-Curtis (Bray and Curtis, 1957) to give consistency in the analytical framework. Firstly, to visualise differences in mammal community composition between social media platforms, Nonmetric Multidimensional Scaling (NMDS) based on Bray-Curtis was employed, using the number of sightings per species per week per platform (45 species * 17 weeks * 5 platforms). This ordination technique allowed a two-dimensional spatial representation of mammal communities to be created in a similar way to Principal Component Analysis (PCA). Unlike PCA, however, non-parametric NMDS used rank orders rather than original values. This was important as the absolute number of social media posts, and thus the expected number of posts for each species even if an identical relative distribution between platforms was assumed, would have skewed results (Zuur et al. 2007). Secondly, to assess statistical significance, a Bray-Curtis Permutational Multivariable Analysis of Variance (PERMANOVA) was used (Anderson, 2014). This compared whether there was a significant difference between the metric centroids summarising the communities recorded on the different platforms. Once the overall model was created, pairwise models were created for post-hoc analysis to understand which specific social media platforms differed. A total of 999 permutations was used to calculate the p-value in all cases. Finally, to understand what species were primarily responsible for differences observed between social media platforms, a Bray-Curtis similarity percentage analysis (SIMPER) was conducted. This involved calculating the mean dissimilarity for each platform, and then disaggregating this value to calculate the contribution percentage of each species to this total (Clarke et al. 2014).

Results

In total, 7,860 unique sightings were posted across the five platforms (Latest Sightings = 3,262; Instagram = 1,800; Facebook =1,698; X = 764; Flickr = 336). Across all platforms combined, 45 species were recorded but species richness recorded varied substantially between platforms (Table 1). Indeed, no one platform recorded all species: Latest Sightings (37 species; 82%), Instagram (32 species; 71%), Facebook (29 species; 64%), X (28 species; 62%); Flickr (32 species; 71%). Just 20 species (44%) were recorded on all five platforms over the 4-month recording period. This included cape buffalo (*Syncerus caffer*), cheetah (*Acinonyx jubatus*), impala (*Aepyceros melampus*), klipspringer (*Oreotragus oreotragus*), and black-backed jackal (*Canis mesomelas*). However, for the majority of species (56%), there were inter-platform differences, even when considering presence-only data (Table 1) rather than relative abundance. Indeed, four species were recorded on only one platform: blesbok (*Damaliscus pygargus phillipsi*) on Facebook; pangolin (*Smutsia temminckii*) and side-striped jackal (*Canis adustus*) on Latest Sightings, and springhare (*Pedetes capensis*) on Flickr. The most common species recorded across all platforms were lion (2,034 sightings; 26%), elephant (1,071 sightings; 14%), and leopard (884 sightings; 11%).

The overall Bray-Curtis dissimilarity for community composition, based on the number of sightings per species per week per platform, was 64.18%. The NMDS plot created to allow visualization of intra- and inter-platform differences in species relative abundance to assess community composition (Fig. 1a) showed that the mammal community recorded via posts on Latest Sightings and Instagram formed tight and discrete clusters. This suggested that the species composition in any given week was very similar to the mean for each of these platforms (low intra-platform variability), but there were substantial differences between platforms (no overlap between clusters). The species composition based on Facebook and X were less tightly clustered (higher intra-platform variability), while the species composition based on Flickr was the most dispersed (highest intra-platform variability). There were partial overlaps in clusters for Facebook, X, and Flickr, suggesting slightly less dissimilarity in community composition between these three platforms relative to one another than between these three platforms relative to Latest Sightings and Instagram.

Unsurprisingly given the visual differences shown in Fig 1a, a one-way PERMANOVA revealed a significant difference in community composition between platforms ($F_{4,16}$ =14.730; P < 0.001). Post hoc pairwise comparisons, Bonferroni-corrected to avoid family-wise error, were significant (P < 0.001) in all cases. SIMPER analysis showed that differential representation of just four of the 45 species between social media platforms was cumulatively responsible for over half (52.84%) of the community dissimilarity: lion = 23.82%; elephant = 11.50%; leopard = 11.00%; spotted hyena (*Crocuta crocuta*) = 6.54%. Lion and leopard were both substantially over-recorded on Latest Sightings and under-recorded on Instagram and Flickr (Fig 1b; 1d). Elephant was substantially over-recorded on Instagram and X (Fig 1c). Spotted hyena was substantially over-recorded on Latest Sightings and under-recorded on X and Flickr (Fig 1d). Although white and black rhinoceroses were not recordable on Latest Sightings because of concerns over poaching, the contribution of these species to overall dissimilarity was relatively low (black = 0.39%; white = 4.50%).

Discussion

We found substantial and significant differences in recorded community composition between social media platforms based on user-generated species sightings, albeit in a temporally and spatially limited study. Despite over-recording key predators such as lion and leopard, the bespoke Latest Sightings app gave the richest and most consistent data. X (formerly Twitter) was the weakest in terms of recorded species community, while Flickr was the least consistent between recording weeks and thus required a longer recording period to get a dataset comparable to other platforms. Facebook and Instagram had intermediate performance for both recorded species richness and consistency between recording weeks.

Reliable monitoring through social media posts relies not only on sufficient people on the ground acting as citizen sensors, but also for records to be uploaded (Tufecki, 2014). Kruger National Park is the most visited National Park in South Africa, with an estimated 1.6 million visitors between April 2022 and March 2023 (van der Merwe, 2023), giving a large number of people potentially able to upload sightings with WiFi connectivity at rest camps and relatively good data coverage (krugerpark.co.za/Maps_of_Kruger_Park-travel/cellphone-map-knp.html), making posts of sightings easier – and thus more likely – than might be the case elsewhere. This means that, as our results show, using sightings posts on social media platforms for monitoring species presence in African field settings is feasible. Importantly, however, our results demonstrate very clearly that it cannot be assumed that different social media platforms generate the same relative species community distribution as one another, nor even the same presence-only species assemblage. This has important applied relevance if social media posts are to be used for iEcology approaches to surveying and monitoring.

As might be expected, charismatic flagship species such as lion and leopard (predators) and elephants (megafauna) (Di Minnin et al., 2013) are over-recorded in relation to their true abundance (37% of sightings were of lion and leopard), while far more abundant and frequently encountered antelope, such as impala and blue wildebeest (*Connochaetes taurinus*), are relatively under-recorded. Lion, leopard and elephant are part of the Big 5 (alongside black rhino and cape buffalo) and are much sought after by wildlife tourists. However, an interesting pattern emerges when drilling deeper into the data. On the image-driven platforms, Instagram and Flickr, elephant is overrepresented while lion and leopard are underrepresented. It seems likely that the relative ease with which elephant can be photographed on a smartphone (Instagram) or using a separate camera (Flickr) compared to lion and leopard, which are often hidden by vegetation, could account for this. A further potential issue with Flickr, which was beyond the scope of this study to investigate, is that users are likely to upload images from a separate camera after (and potentially well after) the sighting time. This behaviour, especially if the lag time is substantial (months or even years), could compromise the value of Flickr for gathering data on species occurrence and temporal change, although studies of Flickr behaviour do not indicate that this lag time is a prominent feature of the platform (e.g. Juhász and Hochmair, 2019)) Despite some charismatic species being over-recorded, users did post sightings of many other mammals, confirming other research that many visitors are not solely motivated by flagship species (Hausmann et al., 2017).

X, especially when it was Twitter, has/had APIs that made it relatively easy to "mine" posts for data. Consequently, it became the most popular platform for passive collection of data (Huang et al., 2022), supporting studies on a wide range of topics including loneliness (Koh and Liew, 2022), disease outbreaks (Jahanbin et al., 2022) and voting decisions (Rita et al., 2023) as well as ecology (Daume, 2016; Hart et al., 2018; Edwards et al., 2022). In this study, however, X only recorded around two-thirds of the full species community across all platforms, such that an exclusive focus on this platform would have resulted in a very incomplete species assemblage. The bespoke Latest Sightings app (which functions in a similar way as a social media platform with posts, likes, and comments) was the best platform as regards species coverage and low variability over time, but was also by far the most active platform, with users specifically posting sightings to assist others. This is reflected in Latest Sightings having the most useable location and behavioural data. With its tightly defined focus and presumably motivated users, this platform is arguably closer to a citizen science interface than a "traditional" social media platform, although it was not set up specifically to gather data. Of the remaining platforms, no one platform dominated, such that a multi-platform approach was greatly superior to use of any one platform in isolation.

For social media platforms to be used for iEcology (Jarić et al., 2021; Pernat et al., 2024) based on wildlife sightings it is vital to understand user motivations and behaviour (Toivonen et al., 2019, Edwards et al., 2021), and to be mindful of the limitations this study highlights. Multiple platforms should be used wherever possible, although in some territories some platforms may have limited or zero uptake. As highlighted by Mangachena et al. (2023), social media posts may also give data on spatial and temporal "hot spots" for some species, which could be helpful for management and tourist marketing. With sufficient posts, social media can even be used to inform conservation decision-making. A study focussing on biodiversity in Bangladesh, for example, found that including Facebook data doubled occurrence records gathered through citizen science, and led to the identification of 4,000-10,000 km^2 additional priority areas (Chowdhury et al., 2024).

In addition, while we show that social media posts can be effective for gathering presence data, the clear over- and under-recording of certain species on different platforms render inferences on population size – and even possibly quantification of relative abundance – speculative at best. Despite these limitations, estimating temporal change in population might be feasible, as shown, for example, for changes in the distribution of Banded Demoiselle (*Calopteryx splendens*) in Britain (O'Neill et al., 2023). This is especially the case when intra-platform consistency is high, as shown in this case for both Latest Sightings and Instagram. Such use of passive data collection might be especially useful in resource-limited conservation contexts (Field et al., 2005; Grantham et al., 2009), particularly in regions where reliable and up-to-date biodiversity data are lacking (Chowdhury et al., 2024).

References

Anderson MJ (2017) Permutational Multivariate Analysis of Variance (PERMANOVA). In Wiley StatsRef: Statistics Reference Online (eds N. Balakrishnan, T. Colton, B. Everitt, W. Piegorsch, F. Ruggeri and J.L. Teugels). https://doi.org/10.1002/9781118445112.stat07841

Berryman AJ, Kirwan GM (2021) Is the tiny hawk (*Accipiter superciliosus*) really a specialized predator on hummingbirds? Using citizen science data to elucidate dietary preferences of a little-known Neotropical raptor. J. Raptor Res 55: 276-280. https://doi.org/10.3356/0892-1016-55.2.276

Bray JR, Curtis JT (1957) An ordination of the upland forest communities of southern Wisconsin. Ecol Monogr 27: 326-349.

Breckheimer IK, Theobold EJ, Cristea NC, Wilson AK, Lundquist JD, Rochefort RM, HilleRisLambers J (2020) Crowd-sourced data reveal social–ecological mismatches in phenology driven by climate. Front Ecol Environ 18: 76–82. https://doi.org/10.1002/fee.2142

Chhaya V, Lahiri S, Jagan MA, Mohan R, Pathaw NA, Krishnan A (2021) Community bioacoustics: studying acoustic community structure for ecological and conservation insights. Front Ecol Evol 9: 706445. https://doi.org/10.3389/fevo.2021.706445

Chowdhury S, Aich U, Rokonuzzaman M, Alam S, Das P, Siddika A, Ahmed S, Labi MM, Marco MD, Fuller RA, Callaghan CT (2023) Increasing biodiversity knowledge through social media: A case study from tropical Bangladesh. BioSci 73: pp.453-459.

Chowdhury S, Fuller RA, Ahmed S, Alam S, Callaghan CT, Das P, Correia RA, Di Marco M, Di Minin E, Jarić I, Labi MM (2024) Using social media records to inform conservation planning. *Cons Bio* 38: p.e14161.

Clarke KR, Chapman MG, Somerfield PJ, Needham HR (2006) Dispersion-based weighting of species counts in assemblage analyses. Mar Eco Prog Ser 320: 11-27.

Daume S (2016) Mining Twitter to monitor invasive alien species – an analytical framework and sample information topologies. Ecol Inform 31: 70–82. https://doi.org/10.1016/j.ecoinf.2015.11.014

di Minin E, Fraser I, Slotow R, MacMillan DC (2013) Understanding heterogeneous preference of tourists for big game species: implications for conservation and management. Anim Conserv 16: 249-258. https://doi.org/10.1111/j.1469-1795.2012.00595.x

di Minin E, Tenkanen H, Toivonen T (2015) Prospects and challenges for social media data in conservation science. Environ Sci 3: 63 https://doi.org/10.3389/fenvs.2015.00063

Edwards T, Jones C, Perkins S, Corcoran P (2021) Passive citizen science: The role of social media in wildlife observations. PLoS One 16: e0255416 https://doi.org/10.1371/journal.pone.0255416

Edwards T, Jones CB, Corcoran P (2022) Identifying wildlife observations on Twitter. Ecol Inform 67: p.101500. https://doi.org/10.1016/j.ecoinf.2021.101500

Field S, Tyre A, Possingham H (2005) Optimizing allocation of monitoring effort under economic and observational constraints. J Wildlife Manage 69: 473-482. https://doi.org/10.2193/0022-541X(2005)069[0473:OAOMEU]2.0.CO;2

Ghermandi A, Sinclair M (2019) Passive crowdsourcing of social media in environmental research: a systematic map. Global Environ 55: 36-47. https://doi.org/10.1016/j.gloenvcha.2019.02.003

Goodenough A, Hart AG (2017) Applied Ecology: Monitoring, managing, and conserving. Oxford University Press: UK.

Grantham H, Wilson K, Moilanen A, Revelo T, Possingham H (2009) Delaying conservation actions for improved knowledge: how long should we wait? Ecol Lett 12: 293-301. https://doi.org/10.1111/j.1461-0248.2009.01287.x

Hammer Ø, Harper D, Ryan P (2001) PAST: Paleontological statistics software package for education and data analysis. Palaeontol Electron 4: 1-9.

Hart AG, Carpenter W, Hlustik-Smith E, Reed M, Goodenough A (2018) Testing the potential of Twitter mining methods for data acquisition: Evaluating novel opportunities for ecological research in multiple taxa. Methods Ecol Evol 9: 2194-2205. https://doi.org/10.1111/2041-210X.13063

Hart AG, Dawson M, Fourie R, MacTavish L, Goodenough A (2022) Comparing the effectiveness of camera trapping, driven transects and ad hoc records for surveying nocturnal mammals against a known species assemblage. Comm Ecol 23: 27-39. https://doi.org/10.1007/s42974-021-00070-7

Hausmann A, Toivonen T, Heikinheimo V, Tenkanen H, Slotow R, Minin ED (2017) Social media reveal that charismatic species are not the main attractor of ecotourists to sub-Saharan protected areas. Sci Rep 7: 1-9. https://doi.org/10.1038/s41598-017-00858-6

Hausmann A, Toivonen T, Fink C, Heikinheimo, V., Tenkanen, H., Butchart, S.H., Brooks, T.M. and Di Minin, E., 2019. Assessing global popularity and threats to Important Bird and Biodiversity Areas using social media data. *Science of the Total Environment*, *683*, pp.617-623.

Huang X, Wang S, Zhang M, Hu T, Hohl A, She B, Li Z (2022) Social media mining under the COVID-19 context: Progress, challenges, and opportunities. Int J Appl Earth Obs Geoinformation 113: 102967. https://doi.org/10.1016/j.jag.2022.102967

Jahanbin K, Jokar M, Rahmanian V (2022) Using Twitter and web news mining to predict the monkeypox outbreak. Asian Pac J Trop Med 15: 236-238. https://doi.org/10.4103/1995-7645.346083

Jarić I, et al. (2021) Invasion Culturomics and iEcology. Conserv Biol 35: 447-451. https://doi.org/10.1111/cobi.13707

Juhász L, Hochmair HH (2019) Comparing the spatial and temporal activity patterns between Snapchat, Twitter and Flickr in Florida. GIS Center. 73.

Koh JX, Liew TM (2022) How loneliness is talked about in social media during COVID-19 pandemic: Text mining of 4,492 Twitter feeds. J Psychiatr Res 145: 317-324. https://doi.org/10.1016/j.jpsychires.2020.11.015

Laor T (2022) My social network: Group differences in frequency of use, active use, and interactive use on Facebook, Instagram and Twitter. Technol Soc 68: 101922. https://doi.org/10.1016/j.techsoc.2022.101922

Lindsey P, Baghai M, Bigurube G, Cunliffe S, Dickman A, Fitzgerald K, Robson A (2021) Attracting investment for Africa's protected areas by creating enabling environments for collaborative management partnerships. Biol Conserv 255: 108979. https://doi.org/10.1016/j.biocon.2021.108979

Mangachena JR, Geerts S, Pickering CM (2023) Spatial and temporal patterns in wildlife tourism encounters and how people feel about them based on social media data from South Africa. J Outdoor Recreat Tour 44: 100642. https://doi.org/10.1016/j.jort.2023.100642

Masciantonio A, Bourguignon D, Bouchat P, Balty M, Rimé B (2021) Don't put all social network sites in one basket: Facebook, Instagram, Twitter, TikTok, and their relations with well-being during the COVID-19 pandemic. PLoS One 16: e0248384. https://doi.org/10.1371/journal.pone.0248384

McKinley D, et al. (2017) Citizen science can improve conservation science, natural resource management, and environmental protection. Biol Conserv 208: 15-28. https://doi.org/10.1016/j.biocon.2016.05.015

Mittermeier JC, et al. (2019) A season for all things: phenological imprints in Wikipedia usage and their relevance to conservation. PLoS Biol 17: e3000146. https://doi.org/10.1371/journal.pbio.3000146

Mitterwallner V, Peters A, Edelhoff H, Mathes G, Nguyen H, Peters W, Heurich M, Steinbauer MJ (2024) Automated visitor and wildlife monitoring with camera traps and machine learning. Remote Sens Ecol Conserv 10: 236-247. https://doi.org/10.1002/rse2.367

Naude VN, Smyth LK, Weidemanm EA, Krochuk BA, Amar A (2019) Using web-sourced photography to explore the diet of a declining African raptor, the Martial Eagle (*Polemaetus bellicosus*). Condor 121: 1-9. https://doi.org/10.1093/condor/duy015

O'Neill D, Häkkinen H, Neumann J, Shaffrey L, Cheffings C, Norris K, Pettorelli N (2023) Investigating the potential of social media and citizen science data to track changes in species' distributions. *Ecol Evol* 13: p.e10063.

Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Syst Rev 10: 89. https://doi.org/10.1186/s13643-021-01626-4

Pernat N, et al. (2024) Overcoming biodiversity blindness: Secondary data in primary citizen science observations. Ecol Solut Evid 5: e12295. https://doi.org/10.1002/2688-8319.12295

Rita P, António N, Afonso AP (2023) Social media discourse and voting decisions influence: sentiment analysis in tweets during an electoral period. Soc Netw Anal Min 13: 46. https://doi.org/10.1007/s13278-023-01048-1

Salas-Picazo RI, Ramírez-Bravo OE, Meza-Padilla I, Camargo-Rivera EE (2023) The role of social media groups on illegal wildlife trade in four Mexican states: A year-long assessment. Global Ecol Conserv 45: p.e02539.

Stafford R, Hart AG, Collins L, Kirkhope CL, Williams RL, Rees SG, Lloyd JR, Goodenough AE (2010) Eu-social science: the role of internet social networks in the collection of bee biodiversity data. PLoS One 5: e14381. https://doi.org/10.1371/journal.pone.0014381

Sumner S, Bevan P, Hart AG, Isaac NJ (2019) Mapping species distributions in 2 weeks using citizen science. Insect Conserv Divers 12: 382-388. https://doi.org/10.1111/icad.12345

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. Biol Conserv 233: 298-315. https://doi.org/10.1016/j.biocon.2019.01.023

Tufekci Z (2014) Big questions for social media big data: Representativeness, validity and other methodological pitfalls. Proc Int AAAI Conf Weblogs Soc Media 8: 505-514.

van der Merwe P (2023) Tourists' feelings and behaviours in crowded areas of the Kruger National Park's southern section. Koedoe 65: 1762. https://doi.org/10.4102/koedoe.v65i1.1762

Zuur AF, Leno EN, Smith GM (2007) Analysing Ecological data. Springer, New York.

Table 1: Mammal community (presence/absence) as recorded on a bespoke app (Latest Sightings; grey) and four general social media platforms (white) (total posts used in the analysis are given under each platform name). Posts about white and black rhinoceros are not permitted on the Latest Sightings app due to poaching concerns and are thus denoted (-).

Common Name	**Scientific name**	**L. Sightings (3,262)**	**Instagram (1,800)**	**Facebook (1,698)**	**X/Twitter (764)**	**Flickr (336)**
African Elephant	*Loxodonta africana*	✓	✓	✓	✓	✓
African Wild Cat	*Felis lybica*	✓	-	✓	✓	-
African Wild Dog	*Lycaon pictus*	✓	✓	✓	✓	✓
Baboon (Chacma)	*Papio ursinus*	✓	✓	✓	✓	✓
Banded Mongoose	*Mungos mungo*	✓	✓	-	-	✓
Black Rhinoceros	*Diceros bicornis*	(-)	✓	✓	✓	✓
Black-Backed Jackal	*Canis mesomelas*	✓	✓	✓	✓	✓
Blesbok	*Damaliscus pygargus phillipsi*	-	-	✓	-	-
Brown Hyena	*Parahyaena brunnea*	✓	✓	-	-	-
Bushbaby	*Galago* spp.	✓	-	-	✓	-
Bushbuck	*Tragelaphus scriptus*	-	✓	✓	✓	✓
Cape Buffalo	*Syncerus caffer*	✓	✓	✓	✓	✓
Caracal	*Caracal caracal*	✓	-	-	✓	-
Cheetah	*Acinonyx jubatus*	✓	✓	✓	✓	✓
Civet	*Civettictis civetta*	✓	✓	-	-	-
Duiker	*Cephalophinae*	✓	-	✓	-	✓
Dwarf Mongoose	*Helogale parvula*	✓	✓	-	-	✓
Eland	*Taurotragus oryx*	✓	✓	-	-	-
Genet	*Genetta* spp.	-	✓	✓	-	✓
Giraffe	*Giraffa camelopardalis*	✓	✓	✓	✓	✓
Greater Kudu	*Tragelaphus strepsiceros*	✓	✓	✓	✓	✓
Hippopotamus	*Hippopotamus amphibius*	✓	✓	✓	✓	✓
Honey Badger	*Mellivora capensis*	✓	✓	-	-	✓
Impala	*Aepyceros melampus*	✓	✓	✓	✓	✓
Klipspringer	*Oreotragus oreotragus*	✓	✓	✓	✓	-
Leopard	*Panthera pardus*	✓	✓	✓	✓	✓
Lion	*Panthera leo*	✓	✓	✓	✓	✓
Nyala	*Tragelaphus angasii*	✓	✓	✓	✓	✓
Pangolin	*Smutsia temminckii*	✓	-	-	-	-
Reedbuck	*Redunca* spp.	✓	✓	-	-	-
Sable	*Hippotragus niger*	✓	-	-	✓	-
Serval	*Leptailurus serval*	✓	-	-	✓	-
Side-Striped Jackal	*Canis adustus*	✓	-	-	-	-
Slender Mongoose	*Galerella sanguinea*	-	✓	-	-	✓
Spotted Hyena	*Crocuta crocuta*	✓	✓	✓	✓	✓
Springbok	*Antidorcas marsupialis*	-	-	✓	-	✓
Springhare	*Pedetes capensis*	-	-	-	-	✓
Steenbok	*Raphicerus campestris*	✓	-	✓	-	✓
Tsessebe	*Damaliscus lunatus*	✓	-	-	-	✓
Vervet Monkey	*Chlorocebus pygerythrus*	✓	✓	✓	✓	✓
Warthog	*Phacochoerus africanus*	✓	✓	✓	✓	✓
Waterbuck	*Kobus ellipsiprymnus*	✓	✓	✓	✓	✓
White Rhinoceros	*Ceratotherium simum*	(-)	✓	✓	✓	✓

Wildebeest	*Connochaetes* spp.	✓	✓	✓	✓	✓
Zebra (Plains)	*Equus zebra*	✓	✓	✓	✓	✓

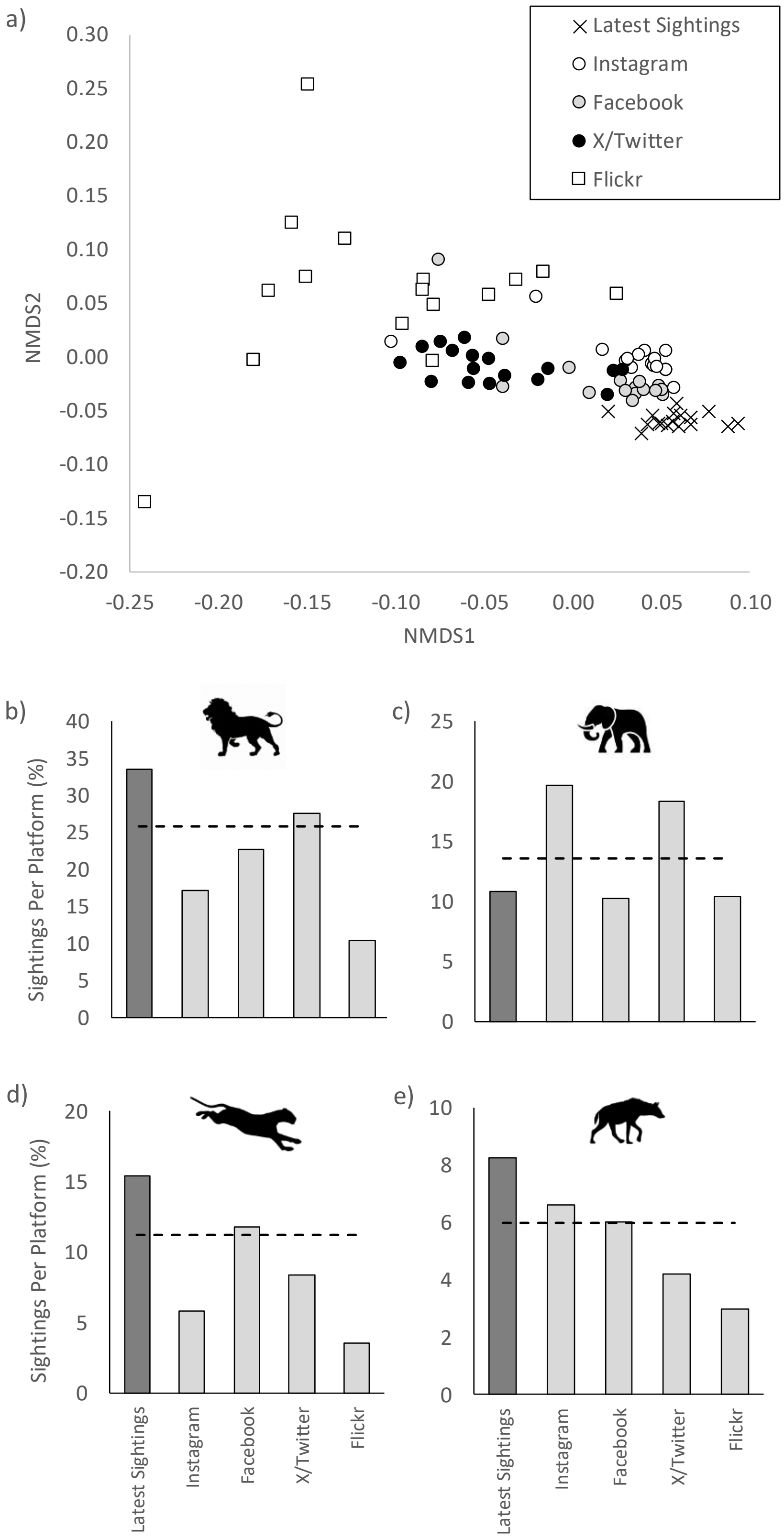

Figure 1: Differences in mammal community in Kruger National Park, South Africa, recorded on five social media platforms showing: (a) intra- and inter-platform differences in community composition (sightings per species per week per platform) using Nonmetric Multi-Dimensional Scaling; and (b-d) the percentage of sightings per platform for the four species that differed most between platforms; lion,

elephant, leopard and spotted hyena (in descending order). On the histograms, the dark grey bar represents the bespoke Latest Sightings app and the light grey bars represent four general social media platforms; the dashed lines show the species-specific average across all platforms.