Data Descriptor

Preprint: Data from a Systematic Review of Crowdsourcing of Research Data from Wildlife Tourism Photographs and Comments Shared on Social Networking Sites

Greg D. Simpson 1,2,* - https://orcid.org/0000-0003-4926-5491
Obelia Walker 1,2,† - https://orcid.org/0000-0001-8567-786X

1 College of Science, Health, Engineering, and Education-Environmental and Conservation Sciences, Murdoch University, Perth 6150, WA, Australia
2 Sukau Ecotourism Research Center (SERC), BEST Society, Lot 1, Pusat Perindustrian, Kolombong Jaya, Jalan Kolombong, 88450 Kota Kinabalu, Sabah, Malaysia.
† Co-First Authors
* Correspondence G.Simpson@Murdoch.edu.au

Abstract: This data descriptor summarizes the process applied and data gathered from 50 publications/papers reporting on the use of photography generated by tourists, tour operators and members of the public, with a particular focus on the crowdsourcing of photographs through online platforms and social networking sites (SNSs) as a method of research for wildlife conservation and ecotourism. The papers were collected in a systematic literature review to inform a pilot study of the feasibility of using SNSs to crowdsource georeferenced photographs of endangered Bornean Pygmy Elephants (Elephas maximus borneensis) taken by ecotourists along the Lower Kinabatangan River region of Sabah, Malaysia. Papers were sourced using the Murdoch University Findit online search tool to search over 100 databases, including Proquest, Scopus and Web of Science. The criteria for a paper to be included in the review (and shared via the dataset attached to this this data descriptor) were that it was peer-reviewed, published in English, between 1997 and the 31 December 2017, had the full text accessible online and reported on a study or studies that utilized photographs that tourists, tour operators and/or members of the public generated and shared via SNSs or online platforms.

Dataset: Dataset submitted for publication as a supplement to this Data Descriptor.

Dataset License: CC-BY

Keywords: Citizen Science; Crowdsourced Data; Ecotourism; Natural Resource Management, Social Media; Photo-elicitation; Photovoice; Wildlife Conservation; Wildlife Tourism

© 2019 by the author(s). Distributed under a Creative Commons CC BY license.
1. Summary

Wildlife tourism is a global industry that has evolved from the inherent human desire to see and interact with wildlife in natural environments [1-5]. Wildlife tourism is, generally, regarded as a non-consumptive activity, with human-wildlife interactions ranging from observing and feeding wildlife thorough to photographing free roaming wild animals [7-10]. As the economic and social significance of wildlife tourism and ecotourism continues to grow [5,11-13], understanding the ecological and social interface of these experiences is critical for achieving long-term sustainability [14,15]. Effective management of wildlife tourism experiences should not only consider the views and perceptions of visitors but should also monitor for impacts on the wildlife that are targeted by these experiences [5,14,16-18]. To achieve management of wildlife tourism that is both responsive and adaptive, it is necessary to have contemporaneous information relating both the visitor experience and the conditions under which the human-wildlife interactions occur [5,16,19]. Such data is, however, often limited [5,9].

Traditionally, obtaining information on how tourists interact with wildlife in natural areas has involved methods such as surveys and interviews, which can be laborious, time-consuming and costly [13,15,16,20-23]. The emerging techniques of modern citizen science and crowdsourcing data can provide an effective alternative to traditional, centralized research methods, particularly when resources and funding are limited [18,24]. Over the past two decades, the combination of and almost ubiquitous connection to the internet, the development of smart devices equipped with geographic positioning system (GPS) services and high quality cameras, and the exponential growth in the willingness of people to share personal information online through social media and other social networking sites (SNSs) has expedited wildlife conservation and tourism research based on collecting data in this way [24-27]. The new alternative of using widespread and readily available data uploaded to social media can provide a rapid and cost-effective way to explore nature-based tourist experiences and activities [13,20,21,28,29]. The dataset shared via this data descriptor informed both the systematic quantitative literature review reported by Walker and others [27] and a pilot study that explored the feasibility of using SNSs to crowdsourced georeferenced photographs of endangered Bornean Pygmy Elephants (Elephas maximus borneensis) taken by ecotourists along the Lower Kinabatangan River region of Sabah, Malaysia. For that reason, the primary function of the systematic quantitative literature review reported by Walker and others [27] was to explore how ecotourist-generated photographs, sourced through existing SNSs, might be used in wildlife conservation and tourism research. The techniques and eligibility criteria used to source and select papers included in the shared dataset and the metadata for the information extracted from those papers are reported in the following sections.

2. Data Description

The identification and screening procedures described in the Methods below identified 48 peer-reviewed papers, one paper published in conference proceedings, and one Masters Thesis that were included and analyzed to produce the dataset shared via this data descriptor. Hereafter, those publications are all referred to as papers. The 50 papers included in the systematic review of Walker et al. [27] were analyzed to provide a structured quantitative overview of the published literature. The data extracted from each paper (Table 1) included publication information, type of article (research, review, or a combination of both), research approach, style of study, geographic information, target species, summary of how the study was conducted, who generated the photographs, what online platform/SNS was used to access the photographs, sample size, if geotagged (GPS referenced) crowdsourced data was discussed, a summary of the main findings of each study and any recommendations for additional research.

It is important to note that in a number of papers, the data crowdsourced from online platforms related to study locations that were different to location(s) published in the researcher attributions. Subsequently, for the purposes of the dataset shared via this data descriptor, the reported location
was defined to be the locality in which the research occurred. The metadata for the variables reported in the Excel spreadsheet associated with this data descriptor are detailed in Table 1.

3. Methods

3.1 Systematic Quantitative Literature Review Method

The dataset shared via this data descriptor was generated from a systemic review of the recent literature regarding the application of photographs crowdsourced from tourists, tour operators and/or members of the public in order to research wildlife conservation and ecotourism. In the style of the reviews reported by Patrioni and others [9] and Parker and others [3,4,30], the systematic literature review of Walker et al. [27] utilized a combination of the approaches of Pickering et al. [31] and the Preferred Reporting Systematic Review (PRISMA) guidelines of Moher at al. [32]. Applying this systematic approach in the identification and screening of the literature, the shared dataset provides a comprehensive overview of the current peer reviewed publications in this field of research.

Table 1: Variables extracted from articles included in the systematic review

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description/ List of Categories</th>
<th>Data Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author (YEAR)</td>
<td>APA In-Text Citation</td>
<td>Text</td>
</tr>
<tr>
<td>Year</td>
<td>Year Paper Was Published</td>
<td>Numeric</td>
</tr>
<tr>
<td>Journal</td>
<td>Full Name of Peer Reviewed Journal</td>
<td>Text</td>
</tr>
<tr>
<td>Location of Study (Country)</td>
<td>If Specified in Paper or N/A = Not Applicable</td>
<td>Text</td>
</tr>
<tr>
<td>Location of Study (Continent)</td>
<td>If Specified in Paper or N/A = Not Applicable</td>
<td>Text</td>
</tr>
<tr>
<td>Target Species</td>
<td>If Specified in Paper or N/A = Not Applicable</td>
<td>Text</td>
</tr>
<tr>
<td>Focus of Study</td>
<td>Descriptive Summary of Research Reported</td>
<td>Text</td>
</tr>
<tr>
<td>Type of Research</td>
<td>Ecological, Social or Both</td>
<td>Categorical</td>
</tr>
<tr>
<td>Photographs Taken/Uploaded By</td>
<td>General Public, Tourists, Tour Operators, General Public &amp; Researchers, Tour Operators &amp; Researchers</td>
<td>Categorical</td>
</tr>
<tr>
<td>Online Platform - Dedicated and/or SNS Name of</td>
<td>Wildbook, Wildbook for Whale Sharks, iNaturalist, eBird, SNS &amp; other or N/A = Not Applicable</td>
<td>Text</td>
</tr>
<tr>
<td>Dedicated Platform</td>
<td>Pic4Turtle, TORSOOI, Waarnemingen, Whale shark photo library or N/A = Not Applicable</td>
<td></td>
</tr>
<tr>
<td>Name of SNS</td>
<td>Flickr, Youtube, Instagram, Facebook, Vimeo, TripAdvisor, Twitter, Panoramio, LinkedIn, combinations thereof or N/A = Not Applicable</td>
<td>Text</td>
</tr>
<tr>
<td>Name of Other Platforms</td>
<td>Wikipedia, PPGIS, OpenStreetMap, Google Earth, GBIF, Map of Life, NBN or N/A = Not Applicable</td>
<td>Text</td>
</tr>
<tr>
<td>Method of Research</td>
<td>Quantitative, Qualitative or Both</td>
<td>Categorical</td>
</tr>
<tr>
<td>Research or Review Article</td>
<td>Review, Research or Both</td>
<td>Categorical</td>
</tr>
<tr>
<td>Sample Size (Num. of Photographs Reported)</td>
<td>As Reported or N/A = Not Applicable</td>
<td>Text</td>
</tr>
<tr>
<td>Sample Size (Num. of Photographs Classified)</td>
<td>Less than 50, 50-100, 100-500, 500-1000, 1000-10000, Greater than 10000 or N/A = Not Applicable</td>
<td>Categorical</td>
</tr>
<tr>
<td>Sample Period</td>
<td>As Reported or N/A = Not Applicable</td>
<td>Categorical</td>
</tr>
<tr>
<td>Sample Period</td>
<td>Less than 1 month, 1-6 months, 6 months - 1 year, 1-2 years, 2-5 years, greater than 5 years</td>
<td>Categorical</td>
</tr>
<tr>
<td>Geotagged Photographs</td>
<td>Yes or No</td>
<td>Binary</td>
</tr>
<tr>
<td>Main Findings</td>
<td>Descriptive Summary of Key Findings</td>
<td>Text</td>
</tr>
<tr>
<td>Future Research</td>
<td>Descriptive Summary of Recommended Research or N/A = Not Applicable</td>
<td>Text</td>
</tr>
</tbody>
</table>
3.2 Identification, Screening and Exclusion of Papers

Academic publications were identified using the Murdoch University Findit online-search tool to perform a global search of over 100 databases, which included Scopus, Web of Science. The Findit search tool also provided access to paper made available through the BONUS+ and ArticleReach databases.

The initial search parameters were specified as being peer reviewed material published in the five years period from 1 January 2011 to 31 December 2017 with the full text of the paper available online. Given the nature of the crowdsourced volunteered geographic information as being a relatively new field of research, these criteria captured a high proportion of the most relevant literature. The Findit online tool was first searched in May 2017 and again in February 2018 using the search terms *wildlife* and *tourism* in combination with the terms/phrases *photo* or *social media* using Boolean searches of *(wildlife AND (photo or social media)) and (tourism AND (photo or social media))* . The inclusion of the search term *photo* was useful in not only capturing studies that utilized photographs, but also studies that used photovoice and photo-elicitation approaches (see Simpson and Walker [18] for discussion of those research methods in the context of research utilizing SNSs).

The preliminary search identified 422 papers (Figure 1) that were screened to reject papers that:

1. Were written in a language other than English;
2. Did not incorporate the use of photographs;
3. Incorporated the use of photographs gathered and analyzed exclusively by researchers; or
4. Could not be applied to wildlife tourism or nature-based tourism research.

The screening process identified 22 papers that were relevant to the research focus of the systematic review. The reference lists of these 22 papers were then crosschecked to locate additional and older foundation publications not identified by the electronic database search that were relevant to the focus of the systemic review. Reference lists of these additional papers were similarly crosschecked, until no additional papers were identified, which implies that all the relevant literature had been identified (Figure 1).

Whilst the application of ecotourist generated photography to wildlife tourism research was the primary focus of the systematic review of Walker et al. [27], papers discussing nature-based tourism were also included, due to the strong overlap in those fields of research. Newsome, Moore and Dowling [3,33] highlight the important role that wildlife observation plays in the recreational experience of visitors to natural areas, which validates the decision to include such papers in the systemic review. The decision to reject papers that reported research based on photographs gathered and analyzed by researchers was necessary to exclude the large volume of wildlife studies that utilize techniques such as camera-trapping [34,35]. Such papers were not relevant to the focus of the review of Walker et al. [27], as that research does not specifically incorporate the use of photographs generated by tourists, tour operators or members of the public. In addition, it is a requirement of the photovoice method adopted for the research informed by this review that photographs be participant generated, further supporting the need to exclude papers based on researcher generated photographs [18,36-38].

After the reiterative crosschecking of reference lists (Figure 1), 50 papers were deemed eligible for inclusion in the systematic review and were included for further analysis [27]. Those 50 papers provide the basis for the dataset shared via the Excel® worksheet attached to this data descriptor.
Figure 1. Systematic literature review process that generated the data shared via this data descriptor.

**Supplementary Materials:** Dataset shared in *Official Lit Review (Version 4) 3-4-2018 - Preprint Supplementary Materials.xls*

**Author Contributions:** G.D.S. & O.W. made equal contributions to this article and as such are co-first authors.

**Funding:** This research received no external funding.

**Acknowledgments:** We wish to thank Professor Albert Teo of Borneo Eco Tours, Sukau Ecotourism Research Centre (SERC) and the Universiti Malaysia Sabah for his encouragement and support for the research project that generated this article. We similarly thank Associate Professor David Newsome from Murdoch University for his guidance and support for the research project that generated this article and for his feedback on an early draft of the text that provided the foundation for this article.

**Conflicts of Interest:** The authors declare no conflict of interest.
References


