Cell PRESS

Geographic searching for ecological studies: a new frontier

Jason W. Karl[1], Jeffrey K. Gillan[2], and Jeffrey E. Herrick[1]

[1] US Department of Agriculture Agricultural Research Service, Jornada Experimental Range, New Mexico State University, Las Cruces, NM, USA
[2] New Mexico State University, Las Cruces, NM, USA

Important ecological knowledge contained in published literature is often overlooked by researchers, resource managers, and policymakers in part because most literature search tools are thematically focused and do not explicitly consider the geographic location of the research. Searching for ecological literature thematically and geographically would improve the ease of accessibility of potentially relevant research findings [1,2] that could promote syntheses and meta-analyses [3,4], provide mechanistic understanding to environmental patterns [5], facilitate evaluations of bias in ecological knowledge [6,7], and limit redundancy in conducting new studies.

Presently, location-based searching for literature from most websites, such as Web of Science, Google Scholar, or SCOPUS, is limited to matching place names from the text of an article. Even the CAB Abstracts database (http://www.cabdirect.org) which indexes agricultural and life science literature by general location (e.g., country, state or province) provides only limited place-name geographic searching. However, finding studies by place names yields incomplete and sometimes irrelevant results and prevents searching for sources from environmentally similar, but geographically separated, areas. Location information (i.e., geographic coordinates or study area descriptions) could be harnessed to enhance literature search tools through georeferencing articles [1].

New tools are beginning to enable geographic literature searching. For example, JournalMap (http://www.journalmap.org) is a simple geographic search engine for ecological literature based on study area descriptions. The online data repository Pangaea (http://www.pangaea.de) effectively maps journal articles by the locations of their data. The GLOBE project (http://globe.umbc.edu) supports searching for, and synthesis of, user-contributed georeferenced studies for land-change science. Additionally, several journals published by Elsevier offer the ability to include supplemental geographic data and interactive maps.

All these efforts, however, suffer from the same shortcomings: they do not contain enough content to realize the potential of geographic literature searching, contributions are either voluntary or limited to specific groups, and most importantly, these efforts operate independently of each other. Consequently, the lack of accessible and searchable georeferenced content is currently limiting the potential of geographic literature searching.

Achieving a meaningful spatial representation of ecological knowledge will require changing how and what information is published in the future, as well as confronting the challenge of georeferencing past studies. Going forward, increasing the prevalence of reporting precise study locations in a standardized format (i.e., via geographic coordinates) is paramount. Although most ecology studies identify their study area and a large proportion report geographic coordinates [1,6,8], this information is effectively unusable if buried in the article text in myriad formats. Standards for reporting study areas must be developed and authors required to report their study area locations according to these standards. Existing geographic standards may be a useful starting point (e.g., [9]), but for georeferencing literature, issues of study area type (e.g., point, linear, or areal) and scale (e.g., site, region, or continent) must also be addressed.

Additionally, the publication and peer-review process is not adequate for ensuring reliable geographic coordinates (i.e., no one is checking the coordinates). Publication of erroneous location information is not uncommon [1,7,8]. Accordingly, processes should be implemented in the submission and editing of manuscripts to check the validity of the study location. For example, tools could be built into manuscript submission websites where specific location information can be entered (e.g., coordinates, place-name lookup using a gazetteer, or digitized on a map) or uploaded [8]. The contributor would then be asked to confirm via a map display of the location that it is correct.

Although these changes will improve search capabilities going forward, they do not address the massive amount of published ecological literature. Realizing that the precision of georeferencing already-published works may not be as good as what can be achieved in the future, large-scale efforts should be launched to georeference existing literature to the best extent possible. Tools and best practices developed by the taxonomic community [9,10] may be useful in this effort. Algorithms for parsing coordinates or place names from articles may help automate the process, but some degree of manual georeferencing and validation will be necessary [7]. It may also be possible to crowd-source article georeferencing or location validation through online scientific reference networks, such as Zotero (http://www.zotero.org) or Mendeley (http://www.mendeley.com).

In the short term, two simple behaviors are imperative: (i) authors should take the initiative to include geographic coordinates and descriptive study area information in their

Corresponding author: Karl, J.W. (jkarl@nmsu.edu).

study area descriptions; and (ii) journal editorial boards and publishers should begin discussions on how broad-scale geographic searching can become standard practice in scientific knowledge discovery.

References

1 Karl, J.W. *et al.* (2013) Geo-semantic searching: discovering ecologically-relevant knowledge from published studies. *Bioscience* (in press)
2 Wallis, P.J. *et al.* (2011) Mapping local-scale ecological research to aid management at landscape scales. *Geogr. Res.* 49, 203–216
3 Hughes, T.P. *et al.* (2002) Detecting regional variation using meta-analysis and large-scale sampling: latitudinal patterns in recruitment. *Ecology* 83, 436–451
4 Van Vliet, N. *et al.* (2012) Trends, drivers and impacts of changes in swidden cultivation in tropical forest-agriculture frontiers: a global assessment. *Global Environ. Change* 22, 418–429
5 Jetz, W. *et al.* (2012) Integrating biodiversity distribution knowledge: toward a global map of life. *Trends Ecol. Evol.* 27, 151–159
6 Martin, L.J. *et al.* (2012) Mapping where ecologists work: biases in the global distribution of terrestrial ecological observations. *Front. Ecol. Environ.* 10, 195–201
7 Fisher, R. *et al.* (2011) Global mismatch between research effort and conservation needs of tropical coral reefs. *Conserv. Lett.* 4, 64–72
8 Shapiro, J.T. and Báldi, A. (2012) Lost locations and the (ir)repeatability of ecological studies. *Front. Ecol. Environ.* 10, 235–236
9 Wieczorek, J. *et al.* (2012) Darwin core: an evolving community-developed biodiversity data standard. *PLoS ONE* 7, e29715
10 Chapman, A.D. and Wieczorek, J. (2006) *Guide to Best Practices for Georeferencing*. Global Biodiversity Information Facility

0169-5347/$ – see front matter. Published by Elsevier Ltd.
http://dx.doi.org/10.1016/j.tree.2013.05.001 Trends in Ecology & Evolution, July 2013, Vol. 28, No. 7