



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2015

Archiving primary data: solutions for long-term studies

Mills, James A ; Teplitsky, Céline ; Arroyo, Beatriz ; Charmantier, Anne ; et al ; Griesser, Michael ; et al ; Zedrosser, Andreas

Abstract: The recent trend for journals to require open access to primary data included in publications has been embraced by many biologists, but has caused apprehension amongst researchers engaged in long-term ecological and evolutionary studies. A worldwide survey of 73 principal investigators (PIs) with long-term studies revealed positive attitudes towards sharing data with the agreement or involvement of the PI, and 93% of PIs have historically shared data. Only 8% were in favor of uncontrolled, open access to primary data while 63% expressed serious concern. We present here their viewpoint on an issue that can have non-trivial scientific consequences. We discuss potential costs of public data archiving and provide possible solutions to meet the needs of journals and researchers.

DOI: <https://doi.org/10.1016/j.tree.2015.07.006>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-120573>

Journal Article

Accepted Version



The following work is licensed under a Creative Commons: Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.

Originally published at:

Mills, James A; Teplitsky, Céline; Arroyo, Beatriz; Charmantier, Anne; et al; Griesser, Michael; et al; Zedrosser, Andreas (2015). Archiving primary data: solutions for long-term studies. *Trends in Ecology Evolution*, 30(10):581-589.

DOI: <https://doi.org/10.1016/j.tree.2015.07.006>

Solutions for archiving data in long-term studies - a reply to Whitlock et al.

James A. Mills¹, Céline Teplitsky^{2,3*}, Beatriz Arroyo⁴, Anne Charmantier³, Peter. H. Becker⁵, Tim R. Birkhead⁶, Pierre Bize⁷, Daniel T. Blumstein⁸, Christophe Bonenfant⁹, Stan Boutin¹⁰, Andrey Bushuev¹¹, Emmanuelle Cam¹², Andrew Cockburn¹³, Steeve D. Côté¹⁴, John C. Coulson¹⁵, Francis Daunt¹⁶, Niels J. Dingemanse^{17,18}, Blandine Doligez⁹, Hugh Drummond¹⁹, Richard H. M. Espie²⁰, Marco Festa-Bianchet²¹, Francesca D. Frentiu²², John W. Fitzpatrick²³, Robert W. Furness²⁴, Gilles Gauthier¹⁴, Peter R. Grant²⁵, Michael Griesser²⁶, Lars Gustafsson²⁷, Bengt Hansson²⁸, Michael P. Harris¹⁶, Frédéric Jiguet², Petter Kjellander²⁹, Erkki Korpimäki³⁰, Charles J. Krebs³¹, Luc Lens³², John D.C. Linnell³³, Matthew Low³⁴, Andrew McAdam³⁵, Antoni Margalida³⁶, Juha Merilä³⁷, Anders P. Møller³⁸, Shinichi Nakagawa³⁹, Jan-Åke Nilsson²⁷, Ian C. T. Nisbet⁴⁰, Arie J. van Noordwijk⁴¹, Daniel Oro⁴², Tomas Pärt³⁴, Fanie Pelletier²¹, Jaime Potti⁴³, Benoit Pujol¹², Denis Réale⁴⁴, Robert F. Rockwell⁴⁵, Yan Ropert-Coudert⁴⁶, Alexandre Roulin⁴⁷, Christophe Thébaud¹², James S. Sedinger⁴⁸, Jon E. Swenson⁴⁹, Marcel E. Visser⁴¹, Sarah Wanless¹⁶, David F. Westneat⁵⁰, Alastair J. Wilson⁵¹, Andreas Zedrosser⁵²

*JAM and CT contributed equally

Authors' affiliations

- 1 10527A Skyline Drive, Corning, NY 14830, USA
- 2 CESCO, UMR7204 Sorbonne Universités-MNHN-CNRS-UPMC, CP51, 55 Rue Buffon, 75005 Paris, France
- 3 Centre d'Ecologie Fonctionnelle et Evolutive UMR 5175, Campus CNRS, 1919 Route de Mende, 34293 Montpellier, Cedex 5, France
- 4 Instituto de Investigacion en Recursos Cinegeticos (IREC) (CSIC-UCLM-JCCM), Ronda de Toledo s/n, 13005 Ciudad, Real, Spain
- 5 Institute of Avian Research, "Vogelwarte Helgoland", An der Vogelwarte 21 D26386 Wilhelmshaven, Germany
- 6 Department of Animal and Plant Sciences, The University of Sheffield, UK
- 7 Institute of Biological and Environmental Sciences, University of Aberdeen, UK
- 8 Department of Ecology and Evolutionary Biology, University of California, 621 Young Drive South, Los Angeles, CA 90095-1606, USA
- 9 CNRS, Université Lyon 1, Université de Lyon, UMR 5558, Laboratoire Biométrie et Biologie Évolutive, 43 boulevard du 11 Novembre 1918, F-69622 Villeurbanne, Cedex, France
- 10 Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2E9
- 11 Department of Vertebrate Zoology, Faculty of Biology, Lomonosov Moscow State University, Leninskie Gory 1/12, 119234 Moscow, Russia
- 12 UMR 5174 EDB Laboratoire Évolution et Diversité Biologique, CNRS, ENFA, Université Toulouse 3 Paul Sabatier, 31062 Toulouse Cedex 9, France
- 13 Department of Evolution, Ecology & Genetics, Research School of Biology, The Australian National University, Canberra, ACT, Australia
- 14 Département de biologie & Centre d'études nordiques, Université Laval, 1045 avenue de la Médecine, Quebec (QC), G1V 0A6, Canada
- 15 29 St Mary's Close, Shincliffe, Durham, DH1 2ND, UK
- 16 Centre for Ecology & Hydrology, Bush Estate, Penicuik, EH26 0QB UK
- 17 Behavioural Ecology, Department of Biology, Ludwig-Maximilians University of Munich, Planegg-Martinsried, Germany
- 18 Evolutionary Ecology of Variation Research Group, Max Planck Institute for Ornithology, Seewiesen, Germany
- 19 Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, A.P. 70-275, México D.F. 04510, México
- 20 Technical Resource Branch, Saskatchewan Ministry of Environment, 3211 Albert Street, Regina, Saskatchewan, Canada S4S 5W6

- 21 Département de biologie, Université de Sherbrooke, 2500 Boulevard de L'Université, Sherbrooke, Québec, Canada J1K 2R1
- 22 School of Biomedical Sciences and Institute of Health and Biomedical Innovation, Queensland University of Technology, Kelvin Grove QLD 4059 Australia
- 23 Cornell Lab of Ornithology, 159 Sapsucker Woods Road, Ithaca, NY 14850, USA
- 24 Graham Kerr Building, University of Glasgow, Glasgow G12 8QQ, U.K.
- 25 Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544-1003, USA
- 26 Anthropological Institute and Museum, University of Zürich, Zürich, Switzerland
- 27 Department of Animal Ecology, Evolutionary Biology Center, Uppsala University, Uppsala, Sweden
- 28 Department of Biology, Lund University, Ecology Building, 223 62, Lund, Sweden
- 29 Grimso Wildlife Research Station, Department of Ecology, Swedish University of Agricultural Sciences (SLU) SE-73091, Riddarhyttan, Sweden
- 30 Section of Ecology, Department of Biology, University of Turku, FI-20014 Turku, Finland
- 31 Department of Zoology, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada
- 32 Terrestrial Ecology Unit, Department of Biology, Ghent University, Ledeganckstraat 35, B-9000 Gent, Belgium
- 33 Norwegian Institute for Nature Research, P.O.Box 5685 Sluppen, N-7485 Trondheim, Norway
- 34 Department of Ecology, Swedish University of Agricultural Sciences, 75007 Uppsala, Sweden
- 35 Department of Integrative Biology, University of Guelph, Guelph, Ontario, Canada, N1G 2W1
- 36 Faculty of Life Sciences and Engineering, University of Lleida, E-25198 Lleida, Spain
- 37 Ecological Genetics Research Unit, Department of Biosciences, PO Box 65 (Biocenter 3, Viikinkaari 1), FIN-00014 University of Helsinki, Finland
- 38 Laboratoire Ecologie, Systématique et Evolution, Equipe Diversité, Ecologie et Evolution Microbiennes, Bâtiment 362, 91405 Orsay Cedex, France
- 39 Evolution & Ecology Research Centre and School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, Australia
- 40 I. C. T. Nisbet & Company, 150 Alder Lane, North Falmouth, MA 02556, USA
- 41 Department of Animal Ecology, Netherlands Institute of Ecology (NIOO-KNAW), P.O.Box 50, 6700 AB Wageningen, The Netherlands
- 42 Institut Mediterrani d'Estudis Avançats IMEDEA (CSIC-UIB), Miquel Marques 21, 07190 Esporles, Mallorca, Spain
- 43 Departamento de Ecología Evolutiva, Estación Biológica de Doñana-CSIC, Av. Américo Vespucio s/n, 41092 Seville, Spain
- 44 Département des Sciences Biologiques, Université du Québec A Montréal, CP 8888-succursale centre-ville, Montreal, Québec, H3C3P8
- 45 Vertebrate Zoology, American Museum of Natural History, New York, NY 10024 USA
- 46 Institut Pluridisciplinaire Hubert Curien, CNRS UMR7178, 23 rue Becquerel 67087 Strasbourg, France
- 47 Department of Ecology and Evolution, University of Lausanne, Switzerland.
- 48 Department of Natural Resources and Environmental Science, University of Nevada Reno, Reno NV 89512, USA
- 49 Department of Ecology and Natural Resource Management, Norwegian University of Life Sciences, P.O.Box 5003, NO-1432 Ås, Norway and Norwegian Institute for Nature Research, P.O.Box 5685 Sluppen, N-7485 Trondheim, Norway
- 50 Department of Biology, Center for Ecology, Evolution, and Behavior, University of Kentucky, Lexington, KY, U.S.A
- 51 Centre for Ecology and Conservation, College of Life and Environmental Sciences, University of Exeter, Cornwall Campus, Penryn TR10 9EZ, UK
- 52 Faculty of Arts and Sciences, Department of Environmental & Health Studies, Telemark UniversityCollege, N-3800 Bø i Telemark, Norway

In our recent paper [1], we discussed some potential undesirable consequences of public data archiving (PDA) with specific reference to long term studies and proposed solutions to manage these issues. We reaffirm our commitment to data sharing and collaboration, both of which have been common and fruitful practices supported for many decades by researchers involved in long-term studies. We acknowledge the potential benefits of PDA [e.g. 2], but believe that several potential negative consequences for science have been underestimated [1, see also 3,4]. The objective of our recent paper [1] was to define practices to simultaneously maximize the benefits and minimize potential unwanted consequences of PDA.

Commenting on our paper, several former and current editors of major ecology and evolution journals [5] acknowledge the need to improve data archiving practices to account for the concerns presented in [1]. The fact that editors of several journals were willing to comment on our paper underlines the importance of this issue, and we are keen to continue this dialogue to identify potential solutions. Following our [1] and Roche et al.'s [6] suggestions, Whitlock et al. [5] endorse as good practice longer embargos (5 years) and encourage cooperation or collaboration with data providers. Both steps are major advances as many of the Principal Investigators (PIs) in [1] have been denied longer embargos, and the practice of consulting PIs to ensure that data files are properly interpreted is not a formal policy in any scientific journal.

We welcome these positive developments but underline three concerns, two of which extend beyond the purview of individual journals:

Whitlock et al. [5] mention that current policies “require only that authors make available the data necessary to recreate the analyses and results in the published manuscript”. For an article that includes an analysis based on a pedigree and individual data, or on lifetime reproductive success and potential predictor variables, this requirement involves providing a detailed data base of the breeding performance of individuals and their progeny over decades. The costs of data gathering, including resources beyond monetary ones, are borne by the data providers and their institutions not by those who would use the data; consequently providing such extensive datasets is sustainable if the data are used only to verify the original analysis. Extending an embargo to five years for such data is a good step, but for studies that extend over decades, a longer embargo is warranted, notably to further encourage potential users to contact PIs to get the latest version of the data, and ideally collaborate.

Databases from long-term studies are an evolving infrastructure that underpin numerous publications. New data are added each year, and errors and omissions are corrected regularly. Over time, archives often include various versions of fragmented datasets which (i) could be combined by

others in ways that the data collectors were already doing or planning to do themselves, or (ii) may differ from each other in ways that are likely to lead to misinterpretation of the data. A single journal's PDA policy cannot ensure that data from long-term studies are not misused. It must be a community decision. Some potential solutions include archiving at institutional servers with separate policies for the distribution of data necessary to reproduce previously published analyses and data requests for additional analyses. The additional analyses would require collaboration with the PI.

Finally, journal editors do not control the policies of funding agencies, but their stature in the community can be influential. Whitlock et al. [5] suggest that funders should set standards for openness. However, long-term studies typically involve several grants and multiple funding agencies, sometimes from different countries. Hence, any discrepancy between their policies can lead to potentially insoluble conflict. Institutions that fund a significant proportion of the research, potentially over decades, may also question the value of continued funding if the data are freely available to individuals from other organizations.

We are encouraged by the letter from Whitlock et al. [5], but believe that there are additional issues that need to be addressed. Some of these may be solved by a more explicit and flexible policy on longer embargos, data storage on institutional servers and involvement of the principal investigators in new analyses using the data they produced, through collaboration or reviews. We hope that this important dialogue will continue.

Cited references

- 1 Mills, J.A. *et al.* (2015) Archiving primary data : solutions for long-term studies. *Trends Ecol. Evol.* 30, 581–589
- 2 Moore, A.J. *et al.* (2010) The need for archiving data in evolutionary biology. *J. Evol. Biol.* 23, 659–660
- 3 Katzner, T. (2015) Do open access data policies inhibit innovation? *BioScience* 65, 1037–1038
- 4 Fenichel, E.P. and Skelly, D.K. (2015) Why should data be free; Don't you get what you pay for? *BioScience* 65, 541–542
- 5 Whitlock, M.C. *et al.* A balanced data archiving policy for long-term studies. *Trends Ecol. Evol.*
- 6 Roche, D.G. *et al.* (2014) Troubleshooting public data archiving: Suggestions to increase

participation. *PLoS Biol.* 12, e1001779