

# Archiving Primary Data: Solutions for Long-Term Studies

James A. Mills, **Céline Teplitsky**, Beatriz Arroyo, Anne Charmantier, Peter. H. Becker, Tim R. Birkhead, Pierre Bize, Daniel T. Blumstein, Christophe Bonenfant, Stan Boutin, Andrey Bushuev, Emmanuelle Cam, Andrew Cockburn, Steeve D. Côté, John C. Coulson, Francis Daunt, Niels J. Dingemanse, Blandine Doligez, Hugh Drummond, Richard H.M. Espie, Marco Festa-Bianchet, Francesca Frentiu, John W. Fitzpatrick, Robert W. Furness, Dany Garant, Gilles Gauthier, Peter R. Grant, Michael Griesser, Lars Gustafsson, Bengt Hansson, Michael P. Harris, Frédéric Jiguet, Petter Kjellander, Erkki Korpimäki, Charles J. Krebs, Luc Lens, John D.C. Linnell, Matthew Low, Andrew McAdam, Antoni Margalida, Juha Merilä, Anders P. Møller, Shinichi Nakagawa, Jan-Åke Nilsson, Ian C.T. Nisbet, Arie J. van Noordwijk, Daniel Oro, Tomas Pärt, Fanie Pelletier, Jaime Potti, Benoit Pujol, Denis Réale, Robert F. Rockwell, Yan Ropert-Coudert, Alexandre Roulin, James S. Sedinger, Jon E. Swenson, Christophe Thébaud, Marcel E. Visser, Sarah Wanless, David F. Westneat, Alastair J. Wilson, and Andreas Zedrosser



# Public data archiving

## Archivage des données:

- en détail suffisant pour pouvoir recréer les analyses
- en open access

## À des fins

- de vérification
- d'archivage / préservation
- De nouvelles analyses (valoriser au plus les données)

## Dans des archives permanentes

- Dryad, TreeBase etc.

## Bénéfice

- Citations via data-papers








## Les problèmes

- Un article suggère que 95% des scientifiques sont favorables au PDA
- Contraste entre jeux de données à court et long terme
- Pas d'information sur les jeux de données a long terme



## Sondage

- 
- Auprès des porteurs de projet sur des jeux de données à long terme
  - Réponses de 73/146 contactés
  - 59 bird studies, 13 mammalian studies, and 1 plant study
- 
- 



	Summary of question	In Favor	Against		
<b>Q1</b>	<b>Position regarding open access archiving</b>	<b>26</b>	<b>46</b>	<b>63.89%</b>	
	Practices for data sharing	Yes	No		
<b>Q10-h</b>	<b>Have you collaborated in meta-analyses?</b>	<b>53</b>	<b>13</b>	<b>80.30%</b>	
<b>Q10-i</b>	<b>Have you involved people to broaden skills?</b>	<b>60</b>	<b>11</b>	<b>84.51%</b>	
<b>Q10-j</b>	<b>Have you supplied the data when asked?</b>	<b>63</b>	<b>5</b>	<b>92.65%</b>	
	Funding issues (asked afterwards, lower sample size)	Yes	No		
<b>Q14</b>	<b>Have you worked without funding?</b>	<b>29</b>	<b>15</b>	<b>65.91%</b>	
<b>Q17</b>	<b>Has your study at any time needed to be curtailed?</b>	<b>16</b>	<b>27</b>	<b>37.21%</b>	

# Coûts potentiels pour la science

## Des erreurs

- Manque de connaissance du système

## Activités redondantes

- Test simultané des mêmes idées
- Vérifier les publications issues des données

## Recherche de financements

- Mutualisation des bénéfices mais pas des coûts
- Difficulté de répondre à des AO

## Moins de collaboration

- Déclin de la participation si contraintes

## Formation des étudiants

- Compétition pour les étudiants en thèse

## Moins d'études à long terme

- Producteur / Chapardeur



# Des solutions

## Augmenter les collaborations

- Partage des bénéfices
- Création de sites pour référencer les données

## Augmenter la communication

- Notification par mail de qui charge les données
- Implication dans le reviewing process

Fournir des données sur des bases confidentielles

## Des embargos plus long

- Analogie avec les brevets
- Demande de données à jour

## Serveurs institutionnels

- Données sauvegardées
- Accès modulable
- Accès aux données à jour, non fragmentées



# Réponses des éditeurs

## A Balanced Data Archiving Policy for Long-Term Studies

Michael C. Whitlock

The American Naturalist

Judith L. Bronstein

The American Naturalist

Emilio M. Bruna

Biotropica

Aaron M. Ellison

Ecological Monographs

Charles W. Fox

Functional Ecology

Mark A. McPeck

The American Naturalist

Allen J. Moore

Journal of Evolutionary Biology

Mohamed A.F. Noor

Evolution

Mark D. Rausher

Evolution

Loren H. Rieseberg

Molecular Ecology

Michael G. Ritchie

Journal of Evolutionary Biology

Ruth G. Shaw

Evolution



# Réponses des éditeurs

## A Balanced Data Archiving Policy for Long-Term Studies

Embargos plus long

Encourager la collaboration mais sans obligation

Demander au collecteur des données de reviewer l'article

Encourager les financeurs à reconnaître la valeur des jeux de données a long terme





RESEARCH INTEGRITY

# *Liberating field science samples and data*

Promote reproducibility by moving beyond “available upon request”

*By* Marcia McNutt,<sup>1\*</sup> Kerstin Lehnert,<sup>2</sup> Brooks Hanson,<sup>3</sup> Brian A. Nosek,<sup>4</sup> Aaron M. Ellison,<sup>5</sup> John Leslie King<sup>6</sup>

Such efforts must recognize that motivations for promoting transparency and reproducibility vary by stakeholder. Researchers want to produce knowledge in new directions and to get credit for their contributions. Funders want to see greater value from their investment.

Journals want to facilitate reproducible science. Repositories want to support their communities and streamline data flow.

**FUNDING, PUBLISHING, AND CULTURE CHANGE.** Transparency and reproducibility in scientific research require invest

---

*“All should credit data creators and accelerate recognition of the value of data in the...system.”*

Funding agencies and journals can guide expectations and set requirements, but top-down mandates alone are unlikely to foster needed cultural changes in scientific communities. Research culture prioritizes publications, innovation, and insight, which puts data stewardship and reuse for down

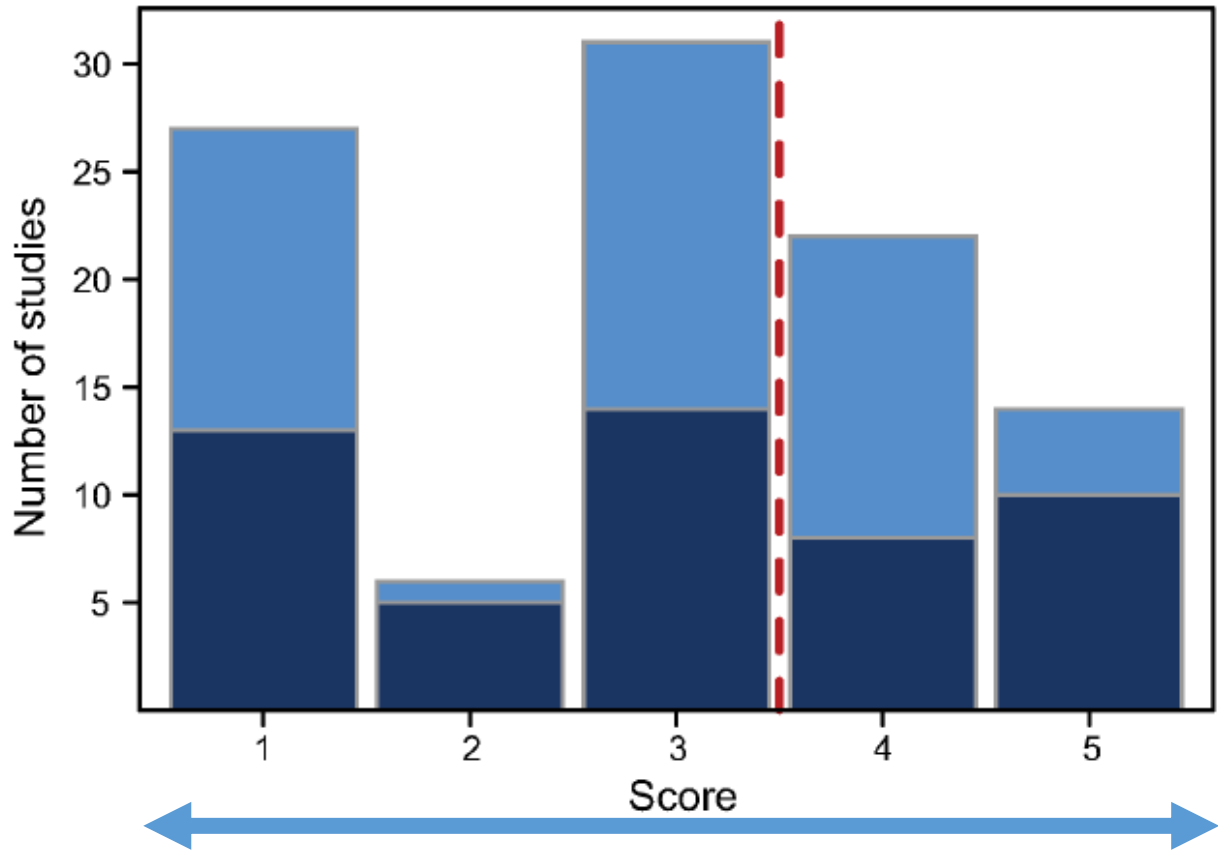
of Dat  
data ci  
publis  
ods. A  
celerat  
the sci  
Des



PERSPECTIVE

# Public Data Archiving in Ecology and Evolution: How Well Are We Doing?

Dominique G. Roche<sup>1,2\*</sup>, Loeske E. B. Kruuk<sup>1,3</sup>, Robert Lanfear<sup>1,4</sup>, Sandra A. Binning<sup>1,2</sup>



Inutilisable

Réutilisable

64% de données archivées dans un format inutilisable



Donner des incentives



# Evolutionary Applications

Evolutionary approaches to environmental, biomedical and socio-economic issues

Original Article

## Intense selective hunting leads to artificial evolution in horn size

Gabriel Pigeon<sup>1,2,\*</sup>, Marco Festa-Bianchet<sup>1</sup>, David W. Coltman<sup>3</sup> and Fanie Pelletier<sup>1,2</sup>

Article first published online: 29 JAN 2016

DOI: 10.1111/eva.12358

© 2016 The Authors. *Evolutionary Applications* published by John Wiley & Sons Ltd.

This is an open access article under the terms of the [Creative Commons Attribution](#) License, which permits use, distribution and reproduction in any

Issue



Evolutionary Applications  
Volume 9, Issue 4, pages  
521–530, April 2016

## Data archiving statement

- The data used in this paper were collected over 39 years and are the infrastructure for several ongoing and planned research programs. They will be available in 10 years from the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.41d7q>. They are also available upon request to anyone who wishes to collaborate with us or repeat our analysis.

REVIEW

Open Access

## Occasional cooperative breeding in birds and the robustness of comparative analyses concerning the evolution of cooperative breeding

Michael Griesser<sup>1,2\*</sup>  and Toshitaka N. Suzuki<sup>2</sup>



## Data sharing

- We adhere to the data sharing policies outlined in Mills et al. [[43](#)].



# Behavioral Ecology

*Behavioral Ecology* will in its turn **promote the recognition of data collectors** by ensuring that **archived data are cited in the bibliography** of the article in which they are used (e.g. [Simmons and Buzatto, 2014](#)). In this way authors will have their data contributions captured by citation metric generators. *Behavioral Ecology* will also insist that re-users likewise **cite both the source of data as well as the original paper** for which those data were generated.

Moreover, along with the submission of any manuscript reporting the use of archived data, authors will be asked to **provide a copy of correspondences between the original data collector and the re-user** that makes it clear that both parties have agreed to the data re-use and that issues surrounding co-authorship have been fully discussed and resolved to the satisfaction of data collector and re-user.





### Mentioned by

- 2 blogs
- 116 tweeters
- 2 Facebook pages

### Readers on

- 133 Mendeley
- 2 CiteULike



**Marco Festa-Bianchet**

@festa\_bianchet

**434**

FOLLOWERS

Several comments on this paper <http://t.co/6uV0qGKoCk> some suggest discussion is warranted, others delighting in scorn, heaping insults

09 Oct 2015

★ Favourite

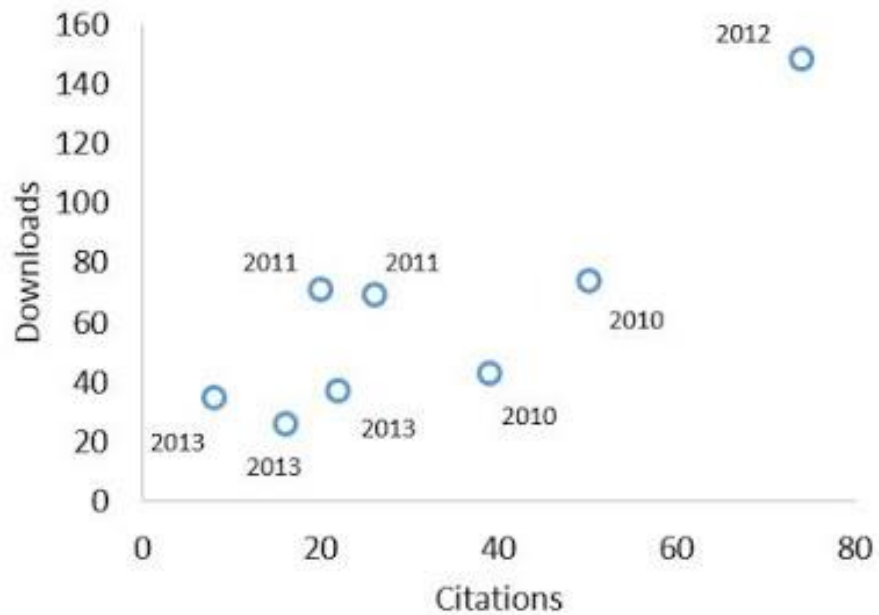


eco-evolutionary  
dynamics

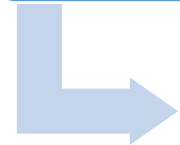


Blog d'Andrew Hendry, McGill University

« Don't worry, be happy. »



16 articles



3788 vues totales



564 téléchargements



0 utilisé dans des publis



eco-evolutionary  
dynamics



Blog d'Andrew Hendry, McGill University

« Don't worry, be happy. »

The simple point is that data will generally be there simply for the asking, regardless of whether it is “archived” online.

One might complain that such data often come with unreasonable demands for co-authorship but, really, if one subscribes to the #OA philosophy about the betterment of society and society, then who cares, really, if you add another author to your paper.

If you want to exclude from co-authorship someone who contributes data to your paper, then surely you shouldn't simultaneously complain when people don't want to share their data.



## Analogie avec des structures expérimentales

- Nécessite beaucoup d'investissement en temps et en argent
- Besoin d'établir des priorités d'accès

## Argument des financements publics à géométrie variable?

- Accès aux structures expérimentales
- Articles non accessibles
- → Text and Data Mining, vote ~ 23/04/2016





**Merci pour votre attention**



Program BioAdapt