Le Plan national pour la science ouverte

Marin Dacos – Conseiller pour la science ouverte
du Directeur général de la recherche et de l'innovation
Ministère de l'enseignement supérieur, de la recherche et de l'innovation - FRANCE

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE

MINISTÈRE
DE L'ENSEIGNEMENT SUPÉRIEUR,
DE LA RECHERCHE
ET DE L'INNOVATION

18/11/2019

Image CC BY Marin Dacos

> *La France s'engage pour que les résultats de la recherche scientifique soient ouverts à tous, chercheurs, entreprises et citoyens, sans entrave, sans délai, sans paiement.*

Pourquoi la science ouverte ?

- Principe : Public money ? Public data !
- Scientifique : une recherche de meilleure qualité.
- Société : une science plus ouverte, plus accessible à la société.

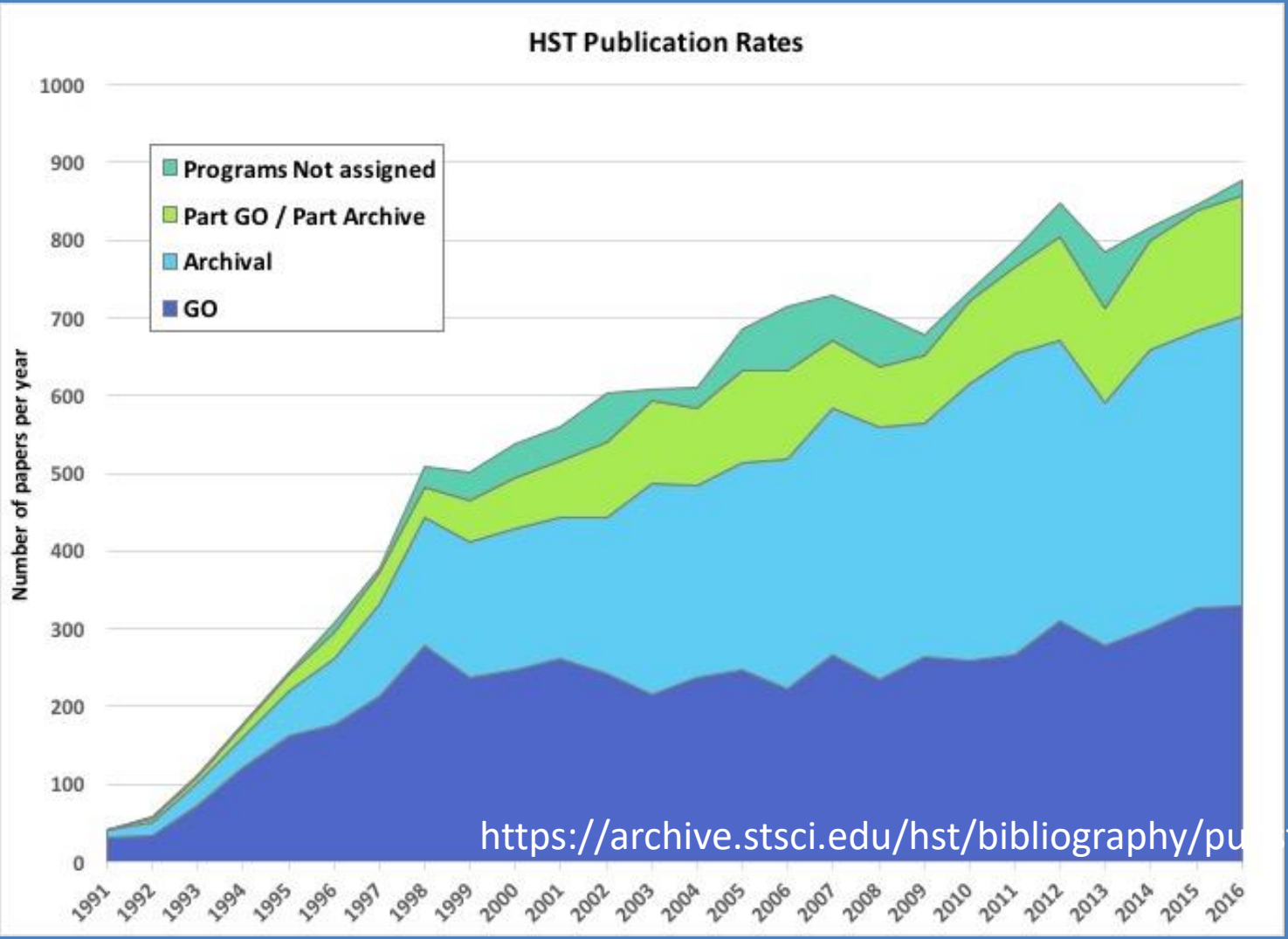
Pourquoi la science ouverte ?

1. Impact : citations et lectures
2. Le “Dernier kilomètre”
3. Efficacité économique
4. Heuristique ! Cumulativité de la science vs redondance des efforts
5. Ethique, intégrité, transparence

« These statistics are based on extensive server log data supplied by Alexandra Elbakyan, the neuroscientist who created Sci-Hub in 2011 as a 22-year-old graduate student in Kazakhstan. I asked her for the data because, in spite of the flurry of polarized opinion pieces, blog posts, and tweets about Sci-Hub and what effect it has on research and academic publishing, some of the most basic questions remain unanswered: **Who are Sci-Hub's users, where are they, and what are they reading?** »

Bohannon, John. « Who's downloading pirated papers? Everyone ». *Science*, 28 avril 2016. doi:10.1126/science.aaf5664.

Cas de l'astronomie : les données « archives » produisent plus de publications que la collecte initiale.



https://archive.stsci.edu/hst/bibliography/pu tat.html

CONTEXTE JURIDIQUE EN FRANCE :
LOI POUR UNE RÉPUBLIQUE NUMÉRIQUE - OCTOBRE 2016

1. Publications : un nouveau droit pour les auteurs

Article 30 : Quand une recherche est financée à au moins 50% par l'Etat, l'auteur dispose de la faculté de déposer dans une archive ouverte le manuscrit auteur accepté, 6 mois (STM) ou 12 mois (SHS) après publication. Si le contrat dit le contraire, il est réputé non écrit.

2. Données : une nouvelle obligation pour les universités et les organismes

Article 6 : principe d'ouverture par défaut pour toutes les données administratives (dans lesquelles on compte les données de la recherche). Cela porte sur des données « achevées ».

PREMIER AXE : **GÉNÉRALISER L'ACCÈS OUVERT AUX PUBLICATIONS**

> *La recherche scientifique est un bien commun que nous devons partager avec tous.*

DEUXIÈME AXE : **STRUCTURER ET OUVRIR LES DONNÉES DE LA RECHERCHE**

●●●

" Les données de la recherche sont la matière première de la connaissance. Les partager, c'est ouvrir de nouvelles perspectives scientifiques . "

Les données, matière première de la connaissance scientifique

Aussi ouvert que possible…

…aussi fermé que nécessaire

Prendre en compte la « longue traîne » (*long tail of data*)

Borgman, Christine L. *Big Data, Little Data, No Data: Scholarship in the Networked World*. Cambridge, Mass.: The MIT Press, 2015, pp.9-10.

Retenir une approche disciplinaire

LA BIBLIOTHÈQUE DES DONNÉES BRÛLE

Current Biology *24*, 94–97, January 6, 2014 ©2014 Elsevier Ltd All rights reserved http://dx.doi.org/10.1016/j.cub.2013.11.014

Report

The Availability of Research Data Declines Rapidly with Article Age

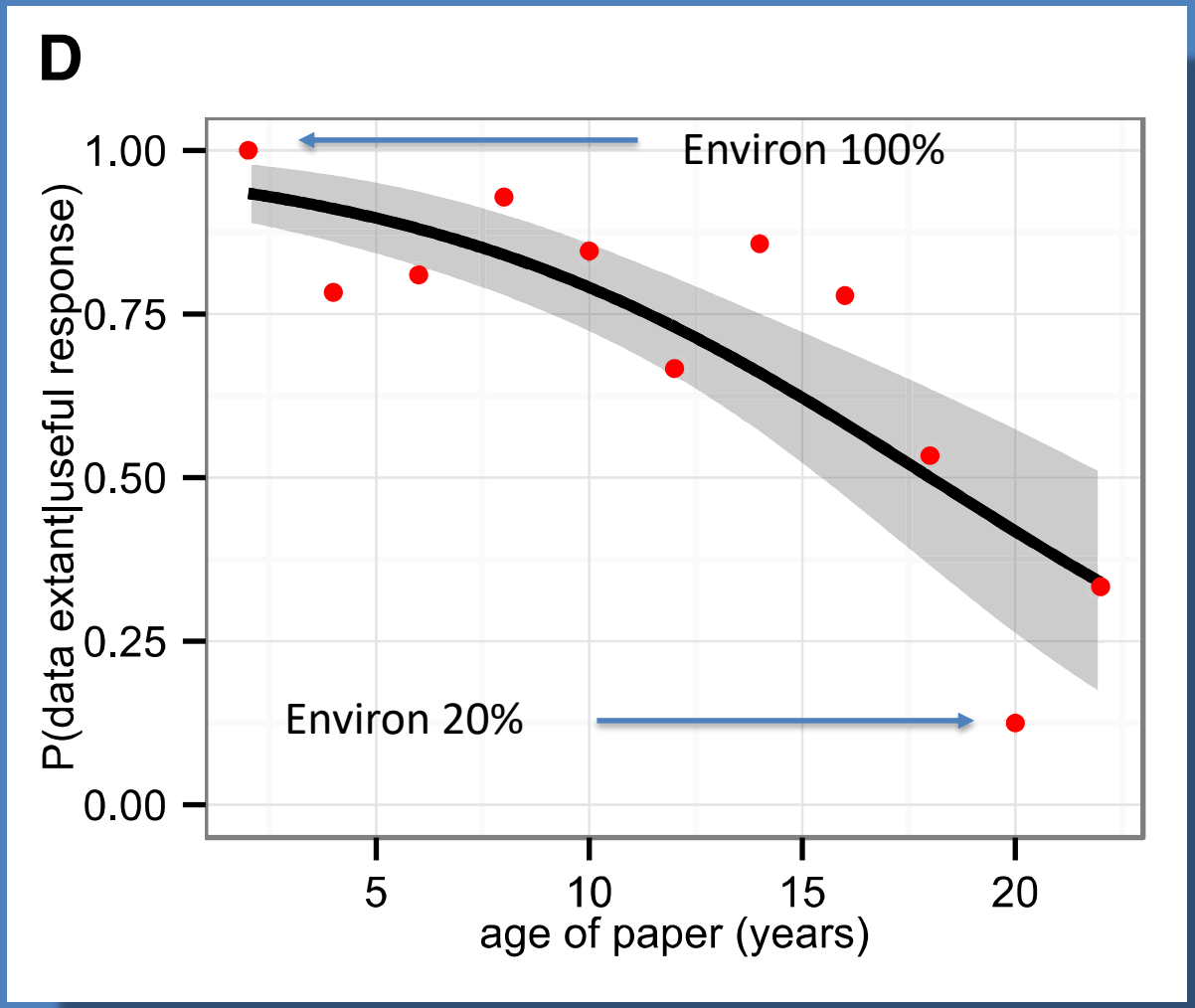
Timothy H. Vines,[1,2,*] Arianne Y.K. Albert,[3] Rose L. Andrew,[1] Florence Débarre,[1,4] Dan G. Bock,[1] Michelle T. Franklin,[1,5] Kimberly J. Gilbert,[1] Jean-Sébastien Moore,[1,6] Sébastien Renaut,[1] and Diana J. Rennison[1]
[1]Biodiversity Research Centre, University of British Columbia, 6270 University Boulevard, Vancouver, BC V6T 1Z4, Canada
[2]Molecular Ecology Editorial Office, 6270 University Boulevard, Vancouver, BC V6T 1Z4, Canada
[3]Women's Health Research Institute, 4500 Oak Street, Vancouver, BC V6H 3N1, Canada
[4]Centre for Ecology & Conservation Biosciences, University of Exeter, Cornwall Campus, Tremough, Penryn TR10 9EZ, UK
[5]Institute for Sustainable Horticulture, Kwantlen Polytechnic University, 12666 72nd Avenue, Surrey, BC V3W 2M8, Canada
[6]Department of Biology, Université Laval, 1030 Avenue de la Médecine, Laval, QC G1V 0A6, Canada

sets (23%) were confirmed as extant. Table 1 provides a breakdown of the data by year.

We used logistic regression to formally investigate the relationships between the age of the paper and (1) the probability that at least one e-mail appeared to work (i.e., did not generate an error message), (2) the conditional probability of a response given that at least one e-mail appeared to work, (3) the conditional probability of getting a response that indicated the status of the data (data lost, data exist but unwilling to share, or data shared) given that a response was received, and, finally, (4) the conditional probability that the data were extant (either "shared" or "exists but unwilling to share") given that an informative response was received.

There was a negative relationship between the age of the paper and the probability of finding at least one apparently working e-mail either in the paper or by searching online (odds ratio [OR] = 0.93 [0.90–0.96, 95% confidence interval

LA BIBLIOTHÈQUE DES DONNÉES BRÛLE



(D) Predicted probability that the data were extant (either ‘‘shared’’ or ‘‘exist but unwilling to share’’) given that we received a useful response. In all panels, the line indicates the predicted probability from the logistic regression, the gray area shows the 95% CI of this estimate, and the red dots indicate the actual proportions from the data.

UK PRIVACY DEBACLES

« Between 24 April 2009 and 29 October 2009 alone, the loss of 878 513 records by 35 organizations is listed. Included in the list is the University of Manchester, after a member of staff emailed an attachment to 469 students with data on 1700 people including information on student disabilities; and Imperial College when six laptops were stolen resulting the loss of medical data containing confidential information on 6000 patients. »

UK Privacy Debacles
https://wiki.openrightsgroup.org/wiki/UK_Privacy_Debacles

« PUBLIC AVAILABILITY OF PUBLISHED RESEARCH DATA IN HIGH-IMPACT JOURNALS » - 2011

OPEN ACCESS Freely available online PLoS one

Public Availability of Published Research Data in High-Impact Journals

Alawi A. Alsheikh-Ali[1,2.], Waqas Qureshi[3.], Mouaz H. Al-Mallah[4,5], John P. A. Ioannidis[2,6,7,8,9]*

1 Institute of Cardiac Sciences, Sheikh Khalifa Medical City, Abu Dhabi, United Arab Emirates, 2 Department of Medicine, Tufts Clinical and Translational Science Institute, Tufts University School of Medicine, Boston, Massachusetts, United States of America, 3 Department of Medicine, Henry Ford Hospital, Detroit, Michigan, United States of America, 4 King Abdul-Aziz Cardiac Center, National Guard Health Affairs, Riyadh, Kingdom of Saudi Arabia, 5 Wayne State University, Detroit, Michigan, United States of America, 6 Stanford Prevention Research Center, Stanford University School of Medicine, Stanford, California, United States of America, 7 Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece, 8 Biomedical Research Institute, Foundation for Research and Technology-Hellas, Ioannina, Greece, 9 Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, United States of America

Abstract

Background: There is increasing interest to make primary data from published research publicly available. We aimed to assess the current status of making research data available in highly-cited journals across the scientific literature.

Methods and Results: We reviewed the first 10 original research papers of 2009 published in the 50 original research journals with the highest impact factor. For each journal we documented the policies related to public availability and sharing of data. Of the 50 journals, 44 (88%) had a statement in their instructions to authors related to public availability and sharing of data. However, there was wide variation in journal requirements, ranging from requiring the sharing of all primary data related to the research to just including a statement in the published manuscript that data can be available on request. Of the 500 assessed papers, 149 (30%) were not subject to any data availability policy. Of the remaining 351 papers that were covered by some data availability policy, 208 papers (59%) did not fully adhere to the data availability instructions of the journals they were published in, most commonly (73%) by not publicly depositing microarray data. The other 143 papers that adhered to the data availability instructions did so by publicly depositing only the specific data type as required, making a statement of willingness to share, or actually sharing all the primary data. Overall, only 47 papers (9%) deposited full primary raw data online. None of the 149 papers not subject to data availability policies made their full primary data publicly available.

Conclusion: A substantial proportion of original research papers published in high-impact journals are either not subject to any data availability policies, or do not adhere to the data availability instructions in their respective journals. This empiric evaluation highlights opportunities for improvement.

Citation: Alsheikh-Ali AA, Qureshi W, Al-Mallah MH, Ioannidis JPA (2011) Public Availability of Published Research Data in High-Impact Journals. PLoS ONE 6(9): e24357. doi:10.1371/journal.pone.0024357

Editor: Isabelle Boutron, University Paris Descartes, France

Received May 17, 2011; Accepted August 6, 2011; Published September 7, 2011

Copyright: ß 2011 Alsheikh-Ali et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: These authors have no support or funding to report.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: jioannid@hsph.harvard.edu

. These authors contributed equally to this work.

« COST OF NOT HAVING FAIR RESEARCH DATA »

European Commission

Cost of not having FAIR research data

Cost-Benefit analysis for FAIR research data

Written by PwC EU Services
March – 2018

Research and Innovation

TROISIÈME AXE : **S'INSCRIRE DANS UNE DYNAMIQUE DURABLE, EUROPÉENNE ET INTERNATIONALE**

" La France s'engage pour que la science ouverte devienne la pratique quotidienne par défaut des chercheurs. "

LE COMITÉ POUR LA SCIENCE OUVERTE

OUVRIR LA SCIENCE !

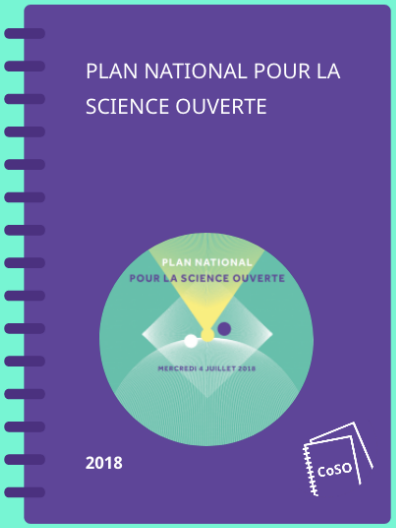
www.ouvrirlascience.fr

OUVRIR
LA SCIENCE !

LA SCIENCE OUVERTE | LE COMITÉ | GROUPES | BLOG | AGENDA | RESSOURCES

Que faut-il faire pour que la science soit plus ouverte ?

DÉCOUVRIR LA SCIENCE OUVERTE



www.ouvrirlascience.fr

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE

MINISTÈRE
DE L'ENSEIGNEMENT SUPÉRIEUR,
DE LA RECHERCHE
ET DE L'INNOVATION

Merci !

Ressources utiles :

https://www.ouvrirlascience.fr/
https://doranum.fr/
https://dmp.opidor.fr/
https://cat.opidor.fr/

> *La France s'engage pour que les résultats de la recherche scientifique soient ouverts à tous, chercheurs, entreprises et citoyens, sans entrave, sans délai, sans paiement.*

#scienceouverte

esr.gouv.fr

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR, DE LA RECHERCHE ET DE L'INNOVATION

Contact :
Marin Dacos
marin.dacos@recherche.gouv.fr
www.ouvrirlascience.fr
Twitter @marindacos

CC BY