## Data sharing, credit and re-use: who's accountable?

Data Management & Open Data; Open Science & Reproducibility Series, Lausanne 2017

> Catriona MacCallum, Advocacy Director, PLOS Member of the Boards OASPA, OpenAire & Royal Society (Publishing); Member of the RCUK OA Practitioner's Group & UUUK Working Group on OA Efficiencies May 2017

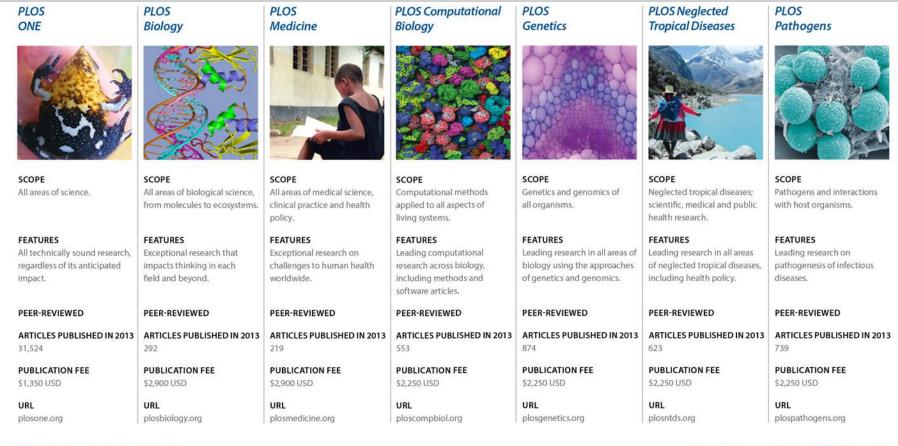
ORCID 0000-0001-9623-2225 @PLOS, @catmacOA

#### **PLOS Mission**

PLOS is a non-profit publisher and advocacy organization with a mission to accelerate progress in science and medicine by leading a transformation in research communication.



#### PLOS – a publisher since 2003



PLOS offers transfers between journals.

Learn more about publishing with PLOS at plos.org.



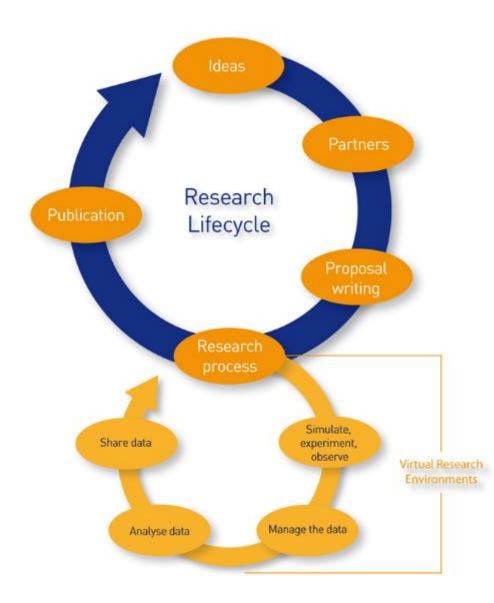
#### What is Open Access ? Free Availability and Unrestricted Use

 Free access – no charge to access
 No embargos – immediately available

✓ Reuse – Creative Commons Attribution License (CC BY) - use with proper attribution







# What is publishing 2

#### © 2014 Jisc: Creative Commons BY-NC-ND

http://www.webarchive.org.uk/wayback/archive/20140615113149/http://www .jisc.ac.uk/whatwedo/campaigns/res3/jischelp.aspx



#### It's no longer just about journals or books



## It's not a cycle...



## ...it's a Network

Image: Andy Lamb, CC BY https://www.flickr.com/photos/speedoflife/8273922515/in/photostream/



#### it's about connections...

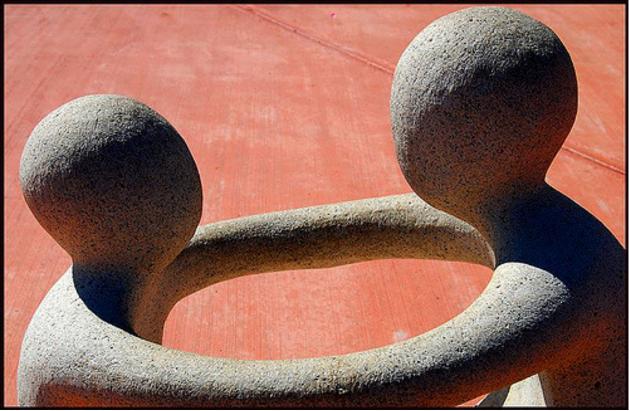


Ingy the Wingy CC BY https://www.flickr.com/photos/ingythewingy/4793928695/in/photostream/

People Organisations Objects, facts, ideas Events



#### and relationships...

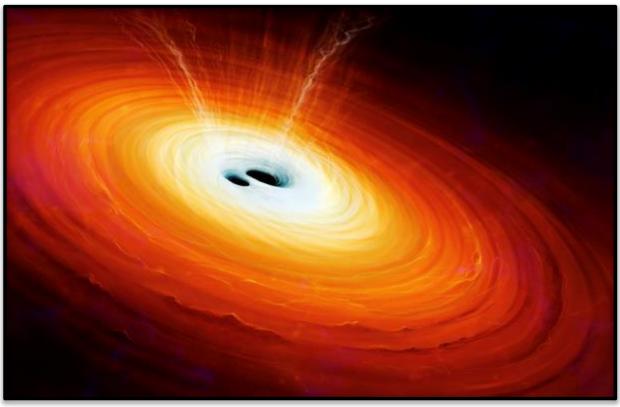


jurek d. Connection CC BY-NC 2.0 https://flic.kr/p/4x8LrS

People Organisations Objects, facts, ideas Events



#### and discovery...



Lwp Kommunikáció, Discovery Science CC BY 2.0 https://flic.kr/p/dyurmR

People Organisations Objects, facts, ideas Events



"Open science is about the way researchers work, collaborate, interact, share resources and disseminate results.

....will bring huge benefits for science itself, as well as for its connection with society. "

Amsterdam Call For Action April 2016 https://english.eu2016.nl/latest/news/2016/04/05/eu-action-plan-for-open-science



## Research INTEGRITY

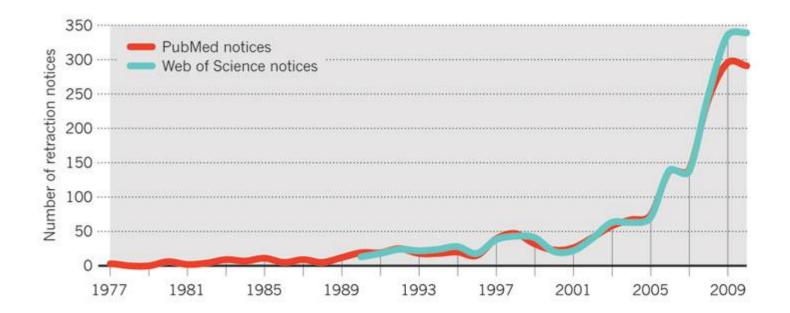


# Public Trust & accountability

Nick Page, Big Ben CC BY 2.0 https://flic.kr/p/k5yH3A

114

#### **Retraction trends**

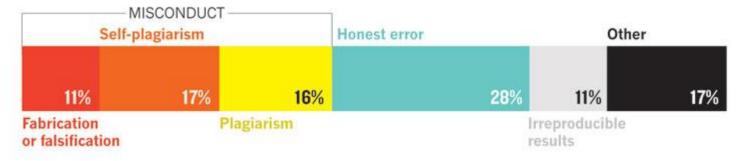


In same period, volume of papers increased by 44%

Van Noorden, Nature 478, 26-28 (2011)



#### Why are papers retracted?



Van Noorden, Nature 478, 26-28 (2011)

# Misconduct accounts for the majority of retracted scientific publications

Ferric C. Fang<sup>a,b,1</sup>, R. Grant Steen<sup>c,1</sup>, and Arturo Casadevall<sup>d,1,2</sup>

Departments of <sup>a</sup>Laboratory Medicine and <sup>b</sup>Microbiology, University of Washington School of Medicine, Seattle, WA 98195; <sup>c</sup>MediCC! Medical Communications Consultants, Chapel Hill, NC 27517; and <sup>d</sup>Department of Microbiology and Immunology, Albert Einstein College of Medicine, Bronx, NY 10461

Edited by Thomas Shenk, Princeton University, Princeton, NJ, and approved September 6, 2012 (received for review July 18, 2012)

A detailed review of all 2,047 biomedical and life-science research articles indexed by PubMed as retracted on May 3, 2012 revealed that only 21.3% of retractions were attributable to error. In contrast, published by the authors of a manuscript in the *Journal of Cell Biology* stated that "In follow-up experiments . . . we have shown that the lack of FOXO1a expression reported in figure 1 is not





#### Is science reliable ?

- Poorly Designed studies
  - small sample sizes, lack of randomisation, blinding and controls
- 'p-hacking' (selective analyses) widespread<sup>1</sup>
- Poorly reported methods & results<sup>2</sup>
- Negative/inconclusive results are not published
- Data not available to scrutinise/replicate

<sup>1</sup>Head ML, Holman L, Lanfear R, Kahn AT, Jennions MD (2015) The Extent and Consequences of P-Hacking in Science. PLoS Biol 13(3): e1002106. doi:10.1371/journal.pbio.1002106 <sup>2</sup>Landis SC, et al. (2012) A call for transparent reporting to optimize the predictive value of preclinical research. Nature 490(7419): 187–191.

**PLOS** 

Science Communication

	AEDICINE		Browse	Publish	About	Search	0
							advanced sea
OPEN ACCESS						13,307	1,792
SAY						Saves	Citations
hy Most Pub			gs Are False			1,448,771 Views	<b>4,859</b> Shares
blished: August 30, 2005	Authors	Metrics	Comments	Related Co	ontent	Download	PDF 🔻
,						Print	Share
in an						Print	Share
lodeling the Framework	Abstract						
lodeling the Framework or False Positive indings	Abstract Summary					CrossMark	S Articles
lodeling the Framework or False Positive indings ias	Summary There is increas	LA 12 12 12	st current published research t			CrossMark Related PLO Most Published Findings Are Fa	S Articles Research Ilse—But a
lodeling the Framework or False Positive indings ias esting by Several	Summary There is increas probability that other studies or	a research claim is tr 1 the same question,	rue may depend on study pow and, importantly, the ratio of t	er and bias, the rue to no relatio	number of nships	CrossMark Related PLO Most Published	S Articles Research Ilse—But a
<b>Ibstract</b> Nodeling the Framework or False Positive indings lias esting by Several ndependent Teams	Summary There is increas probability that other studies or among the relat	a research claim is tr 1 the same question, ionships probed in ea	rue may depend on study pow	er and bias, the rue to no relation ework, a resear	number of nships ch finding is	CrossMark Related PLO Most Published Findings Are Fa Little Replicatio	PS Articles Research Ilse—But a n Goes a
bstract Nodeling the Framework or False Positive indings lias esting by Several idependent Teams corollaries Nost Research Findings	Summary There is increas probability that other studies or among the relat less likely to be smaller; when th there is greater	a research claim is tr n the same question, ionships probed in ea true when the studie here is a greater num flexibility in designs,	rue may depend on study pow and, importantly, the ratio of t ach scientific field. In this fram es conducted in a field are sma ber and lesser preselection of definitions, outcomes, and an	er and bias, the rue to no relation ework, a researd aller; when effec tested relations allytical modes;	number of nships ch finding is t sizes are ships; where when there	CrossMark Related PLO Most Published Findings Are Fa Little Replicatio Long Way	PS Articles Research Ilse—But a n Goes a rotentially n Findings Be
Abstract Modeling the Framework or False Positive Findings Dias esting by Several ndependent Teams Corollaries Most Research Findings are False for Most Research Designs and for	Summary There is increas probability that other studies or among the relat less likely to be smaller; when th there is greater is greater finance scientific field in	a research claim is tr n the same question, ionships probed in ea true when the studie here is a greater num flexibility in designs, ial and other interest n chase of statistical	rue may depend on study pow and, importantly, the ratio of t ach scientific field. In this fram es conducted in a field are sma ber and lesser preselection of definitions, outcomes, and an and prejudice; and when mor significance. Simulations show	er and bias, the rue to no relation ework, a researd aller; when effec tested relations alytical modes; e teams are invo w that for most s	number of nships ch finding is t sizes are ships; where when there olved in a study designs	CrossMark Related PLO Most Published Findings Are Fa Little Replicatio Long Way When Should P False Research Considered Acc When Should P	PS Articles Research Ilse—But a n Goes a otentially n Findings Be ceptable? rotentially
Abstract Nodeling the Framework or False Positive Findings Bias Festing by Several Independent Teams Corollaries Aost Research Findings Are False for Most Research Designs and for Aost Fields Claimed Research Findings May Often Be	Summary There is increas probability that other studies or among the relat less likely to be smaller; when th there is greater is greater finance scientific field in and settings, it	a research claim is tr n the same question, ionships probed in ea true when the studie here is a greater num flexibility in designs, tial and other interest o chase of statistical is more likely for a re	rue may depend on study pow and, importantly, the ratio of t ach scientific field. In this fram es conducted in a field are sma ober and lesser preselection of definitions, outcomes, and an and prejudice; and when mor	er and bias, the rue to no relatio ework, a researd aller; when effec tested relations alytical modes; e teams are invo w that for most s true. Moreover, f	number of nships ch finding is t sizes are ships; where when there olved in a study designs for many	CrossMark Related PLO Most Published Findings Are Fa Little Replicatio Long Way When Should P False Research Considered Acc	PS Articles Research Ilse—But a n Goes a rotentially n Findings Be ceptable? rotentially n Findings Be





- Multi-disciplinary
- Online only
- Open access
- Large, independent editorial board
- Manuscripts assessed only on the rigour of the science, not the novelty/scope of the topic





About

Browse

Search

Q

advanced search

#### The Missing Pieces: A Collection of Negative, Null and Inconclusive Results

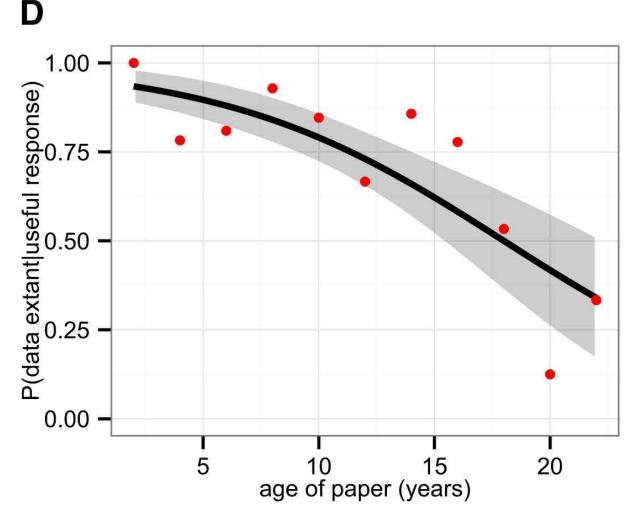
The publication of negative, null and inconclusive results is important to provide scientists with balanced information and avoid the duplication of efforts testing similar hypotheses, which waste valuable time and research resources in the proc...

More >



## Data





#### Data Availability

Probability of finding the data associated with a paper declined by 17% every year

Vines, Timothy et al. "The Availability of Research Data Declines Rapidly with Article Age." Current Biology 24, no. 1 (June 1, 2014): 94–97. doi:10.1016/j.cub.2013.11.014.



### PLOS Data Policy

- PLOS journals require authors to make all data underlying the findings described in their manuscript fully available without restriction, with rare exception.
- When submitting a manuscript online, authors must provide a Data Availability Statement describing compliance with PLOS's policy.

Since March 2014



Making Progress Toward Open Data: Reflections on Data Sharing at PLOS ONE

Data Availability: Biodiversity results, including GIS-ready datasets for open-access use, are available online at http://BiodiversityMapping.org and the Dryad Digital Repository: (http://dx.doi.org/10.5061/dryad.6rv61).

Example Data Availability Statement from Jenkins CN, et al. PLoS ONE. https://doi.org/10.1371/journal.pone.0145064.

Meg Byrne EveryONE May 8 2017: Making Progress Toward open Data http://blogs.plos.org/everyone/2017/05/08/making-progress-towardopen-data/



#### External Data Advisory Group

Academic Chair: Phil Bourne

• 40 experts across the world with representatives from all PLOS journals



#### Guidance for Contributors

### • FAQs consistently updated

http://journals.plos.org/plosone/s/data-availability#loc-faqs-for-data-policy

### • Recommended repositories

http://journals.plos.org/plosone/s/data-availability#loc-recommended-repositories



# What data are required and what is meant by minimal data set?

- PLOS defines the "minimal data set" as the data set used to reach the conclusions drawn in the manuscript with related metadata and methods, and any additional data required to replicate the reported study findings in their entirety:
  - The values behind the means, standard deviations and other measures reported;
  - The values used to build graphs;
  - The points extracted from images for analysis.
- Authors do not need to submit their entire data set, or the raw data collected during an investigation.
  - Just those relevant to the analyses in the paper.



#### Unacceptable Data Access Restrictions

- Authors will not share data because of personal interest (e.g. patents or potential future publications).
- Conclusions depend on proprietary data.
  - data owned by commercial interests
  - copyrighted data that the owners will not share, e.g., data from a pharmaceutical company that will share the data only with regulatory agencies for purposes of drug approval, but not with researchers.



### Internal Checks: PLOS ONE

- At submission: check for unacceptable
   restrictions to access
- During review: Editors & Reviewers assess
   underlying data
- At accept: check statements & ensure clinical datasets have no potentially identifying information
- Post-publication: work with authors as needed



# Possible exceptions to making data publicly available include

- Data cannot be made publicly available for ethical or legal reasons, e.g., public availability would compromise patient confidentiality or participant privacy.
  - Adherence to the PLOS data policy must never breach patient confidentiality.
- Data deposition could present some other threat, such as revealing the locations of fossil deposits, endangered species, or farms/other animal enclosures etc.

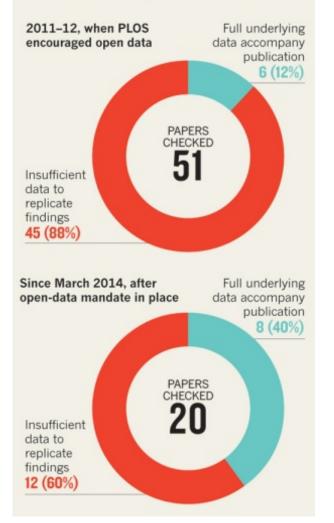


# >65,000 papers published with a data statement at PLOS



#### FREE THE DATA

In 2014, open-access publisher PLOS introduced a requirement that authors who publish in its journals make their underlying data freely available online. An informal audit of one type of population genetics study in one journal, *PLoS ONE*, shows that not everyone is complying — but the mandate is still a boon for the open-data movement.



#### Data Availability: Is it working?

#### 2014: An increase in data sharing<sup>1</sup>:

- from 12% before the policy to 40%
- even up to as much as 76%

#### 2016: Same study<sup>2</sup>

- compliance now 67%

## Not seeing full compliance but we are seeing a MASSIVE improvement

Source:

'1.Confusion over publisher's pioneering open-data rules' *Nature* 515, 478 (27 November 2014) doi:10.1038/515478a
2. Tim Vines, *pers commun (to Meg Byrne, PLOS).*



#### Where are the Data (PLOS ONE)?

In 2016 ~4,000 datasets associated with PLOS articles were deposited in open repositories.

Time	Papers with DAS	Data in Submission Files (#)	Data in Submission Files (%)	Data in Repositories (Estimate)	Data upon Request (Estimate)
Q2-Q4 2014	9491	7918	74%	11%	10%
Q2-Q4 2015	22142	15382	69%	14%	12%
	Dryad	Figshare	NCBI	Github	
Q2-Q4 2014	152	210	551	37	
Q2-Q4 2015	551	753	1229	174	



DAS = Data availability statement

### Data sharing at PLOS ONE

- Very few submissions rejected because of authors' unwillingness or inability to share data (<0.1%)
- Steady growth in publicly available datasets via public data repositories such as the NCBI databases, Figshare or Dryad.
  - ~20% in 2016 low but the growth is encouraging
- 60% of articles include data in the main text and supplementary information
  - supporting information also deposited to Figshare (each item has its own DOI).
- 20% have data available upon request
  - restrictions acceptable under our policy
- Editor & reviewer comments on data availability more frequent
  - from 18% of submissions in 2014 to 24% in 2016 —
  - this is in addition to the yes/no question in the review form asking reviewers to indicate whether the paper complies with the second se

Meg Byrne EveryONE May 8 2017: Making Progress Toward open Data CC BY http://blogs.plos.org/everyone/2017/05/08/making-progress-toward-open-data/

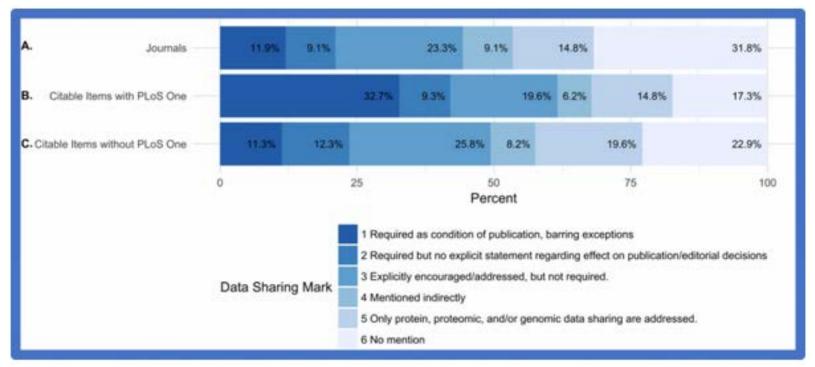
#### Research about data sharing

- <u>PLOS Open Data Collection</u> highlights papers that address issues of data sharing in various scientific disciplines and research showing a correlation between publicly available data and increased impact (for example, citation rates).
- <u>PLOS ONE 10-year Anniversary Datasets Collection</u> highlights specific examples of well-reported or widely used datasets.



### PLOS ONE effect

- A citable item that is open access is much more likely to be published in a journal with a data sharing requirement.
- The proportion of open access journals that require data sharing is much larger than the proportion of subscription journals (64.3% vs 11.3%).
- PLOS ONE significantly increases the proportion of research articles published with a data sharing requirement in biomedical journals



Vasilevsky, Nicole A., Jessica Minnier, Melissa A. Haendel, and Robin E. Champieux. "Reproducible and Reusable Research: Are Journal Data Sharing Policies Meeting the Mark?" PeerJ 5 (April 25, 2017): e3208. doi:10.7717/peerj.3208.



## Challenges

#### QUESTIONS WE DON'T KNOW ANSWERS TO YET

- Treatment of software/code
- How should materials sharing differ
- What to do with big data?
- Do we need better/more aligned consenting for patient studies?
- Best practices for data access committees?
- How to fund data access committees?
- Preservation of obsolete formats?
- How to cite data & credit data reuse?



## The Culture of Evaluation



Incentive	Intended effect	Actual effect
"Researchers rewarded for increased number of publications."	reased number of provide a means of evaluating papers''; poor method false discovery rates le selection of bad scient Mcelreath, 2016); reduced review	
"Researchers rewarded for increased number of citations."	Reward quality work that influences others.	Extended reference lists to inflate citations; reviewers request citation of their work through peer review
"Researchers rewarded for increased grant funding."	"Ensure that research programs are funded, promote growth, generate overhead."	Increased time writing proposals and less time gathering and thinking about data. Overselling positive results and downplay of negative results.
Increase PhD student productivity	Higher school ranking and more prestige of program.	Lower standards and create oversupply of PhDs. Postdocs often required for entry-level academic positions, and PhDs hired for work MS students used to do.
Reduced teaching load for research- active faculty	Necessary to pursue additional competitive grants.	Increased demand for untenured, adjunct faculty to teach classes.
"Teachers rewarded for increased student evaluation scores."	"Improved accountability; ensure customer satisfaction."	Reduced course work, grade inflation.
"Teachers rewarded for increased student test scores."	"Improve teacher effectiveness."	"Teaching to the tests; emphasis on short-term learning."
"Departments rewarded for increasing U.S. News ranking."	"Stronger departments."	Extensive efforts to reverse engineer, game, and cheat rankings.
"Departments rewarded for in- creasing numbers of BS, MS, and PhD degrees granted."	"Promote efficiency; stop students from being trapped in degree programs; impress the state legislature."	"Class sizes increase; entrance requirements" decrease; reduce graduation requirements.
"Departments rewarded for increasing student credit/contact hours (SCH)."	"The university's teaching mission is fulfilled."	"SCH-maximization games are played": duplication of classes, competition for service courses.

#### TABLE 1. GROWING PERVERSE INCENTIVES IN ACADEMIA

Modified from Regehr (pers. comm., 2015) with permission.

"As competition for jobs and promotions increases, the inflated value given to publishing in a small number of socalled "high impact" journals has put pressure on authors to rush into print, cut corners, exaggerate their findings, and overstate the significance of their work.

Such publication practices, abetted by the hypercompetitive grant system and job market, are changing the atmosphere in many laboratories in disturbing ways."

Rescuing US biomedical research from its systemic flaws Bruce Alberts, Marc W. Kirschner, Shirley Tilghman, and Harold Varmus PNAS | April 22, 2014 | vol. 111 | no. 16 | 5773–5777 doi: 10.1073/pnas.1404402111

#### "Career decisions for Early Career Researchers are essentially arbitrary as they are based on so few publications and a hit or miss review process"

'Scholarly publishing: a perspective from an early career academic', COASP 2015, Derek Groen (University College London)





## Now properly co editor who trans bring civility back

The tone used by this reviewer is unacceptably aggressive and accusatory. The reviewer assigns us dark motives when we omit to cite one favoured paper and when we don't provide (in the reviewer's opinion) enough information about the study site. The conclusions drawn by the reviewer about our study site, based on watching youtube videos are frankly ignorant! [...]

If I were the first author of this MS, I probably would not be writing this email. [...] However, the first author of this MS is a graduate student, at the start of her career and her publishing experience, and a review such as this one is incredibly discouraging.

Dear Dr xxxxx,

I am writing regarding the reviews received for MS xxxx. The anger at the comments provided by Reviewer #2. The tone used by this motives when we omit to cite one favoured paper and when we don't proconclusions drawn by the reviewer about our study site, based on wat passages that I believe did not have to be written in such a confront constructive fashion.

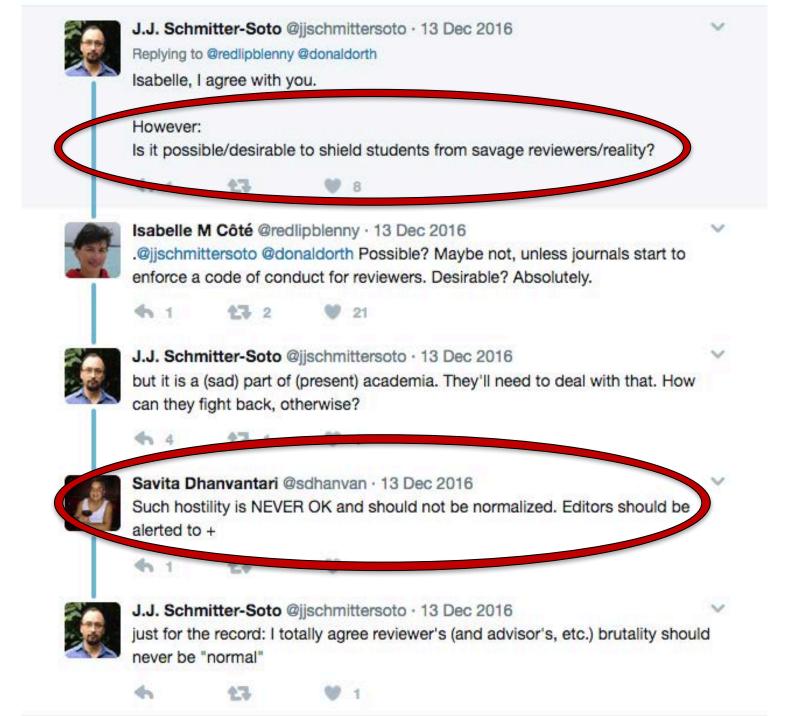
oppressive and accusatory. The reviewer assigns us dark opinion) enough information about the study site. The os (!), are frankly simply ignorant. Eve highlighted in red below the ewers #1 and 3 make many of the same points, but in a much more

If I were the first author on this MS, I probably would not be writing semail. I have received similarly savage reviews before, have read them, been angry, and moved on. However, the first author of this MS is a gradual student, at the start of her career and her publishing experience, and a review such as this one is incredibly discouraging. I would therefore ask you either not to use this reviewer again (and certainly not for papers by graduate students) or exercise your own editorial control and paraphrase when you next transmit this reviewer's comments. I would also really like you to forward this email to the offending reviewer so that he/she realises how their reviews are perceived.

Thank you for your consideration.





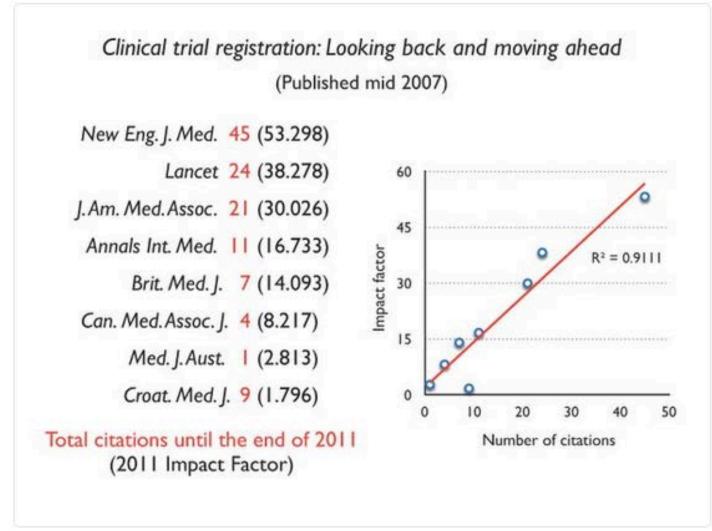


## Current culture embeds status quo

- Researchers gain from publishing in 'designer' journals
- Journals gain financially from their brand/ Journal Impact factor
- Institutions gain financially by hiring and firing based on where researchers publish, not on what they publish (or the mission of the University)
- Research assessment by funders often based on very few publications and brand/impact factor (some are changing)



## Imperfect Impact





#### Impact factors mask huge variation in citations if you use it you are dishonest and statistically illiterate <u>@Stephen\_Curry #COASP</u>

COASP7 'Research and researcher evaluation' (2015), Stephen Curry (Imperial College London) – available soon from OASPA website





HOME	ABO
------	-----

Search

New Results

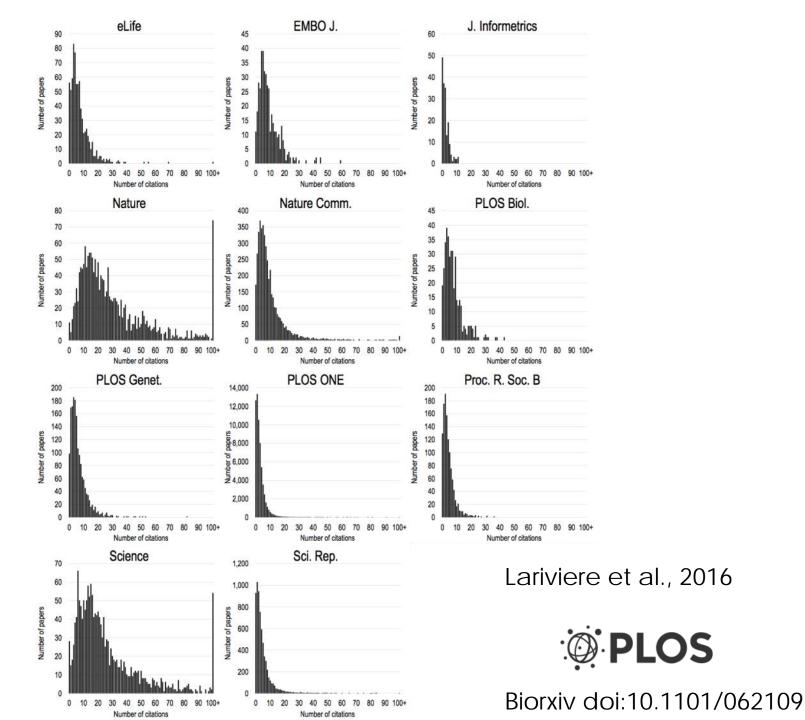
Cold Spring Harbor

#### A simple proposal for the publication of journal citation distributions

			er, 🐵 Catriona J MacCallum, 🐵 Swaminathan, 🐵 Stuart Taylor, (	
doi: http://d	x.doi.org/10.110	/062109		
This article is	a preprint and has	not been peer-r	eviewed [what does this mean?].	
Abstract	Info/History	Metrics	Supplementary material	Preview PDF

#### ARTICLE USAGE

Show by month	Abstract	PDF
Total	22,495	11,381
	Picked up by 7 news outlets	
550	Blogged by 19	
	Tweeted by 554	
	Mentioned by 1 peer review sites	
ee more details	On 10 Facebook pages	
	Mentioned in 1 Google+ posts	
	88 readers on Mendeley	
	5 readers on CiteULike	



## Cultural Change



#### EU COUNCIL CONCLUSIONS ON THE TRANSITION TOWARDS AN OPEN SCIENCE SYSTEM

Removing barriers and fostering incentives (7)

- scientific quality should be based on the work itself
- develop better quality assurance in review and evaluation systems.
- incentives to reward researchers (and research stakeholders) for sharing the results of their research for reuse;
- explore mechanisms to change the ways of doing science.
- collaborate in particular on incentives for an internationally accepted system for data citation



27th May 2016

## Change the Incentives













#### Declaration on Research Assessment

- A worldwide initiative, spearheaded by the ASCB (American Society for Cell Biology), together with scholarly journals and funders
- Focuses on the need to improve the way in which the outputs of scientific research are evaluated:
  - the need to eliminate the use of journal-based metrics, such as Journal Impact Factors, in funding, appointment, and promotion considerations;
  - "need to assess research on its own merits rather than on the basis of the journal in which the research is published"



## Credit: Persistent identifiers and metadata

- Inability to link data to papers & papers to data & papers & data to people
- No separate identifiers for figures, tables, supplementary material etc
- Low adoption of persistent identifiers among Researchers, publishers and data repositories
- Persistent identifiers for Funders & Institutions in flux
  - but being developed





Next-generation metrics:

#### Responsible metrics and evaluation for open science Report of the European Commission Expert Group on Altmetrics March 2017

#### **RECOMMENDATION #8:**

The European research system and Open Science Cloud should adopt ORCID as its preferred system of unique identifiers, and an ORCID iD should be mandatory for all applicants and participants in FP9. Unique identifiers for individuals and research works will gradually improve the robustness of metrics and reduce administrative burden. ORCID provides researchers with a unique ID and associates this ID with a regularly updated list of publications. It is already backed by a growing number of funders across Europe (<u>http://about.orcid.org/</u>). The EC and ERC should utilise ORCID IDs for grant applications, management and reporting platforms, and the benefits of ORCID need to be better communicated to researchers and other stakeholders (Galsworthy & McKee, 2013).



Integrating ORCID iDs in publishing workflows





PLOS ONE

TENTH ANNIVERSARY

 LOGIN • HELP • REGISTER • UPDATE MY INFORMATION • JOURNAL OVERVIEW MAIN MENU . CONTACT US . SUBMIT A MANUSCRIPT . INSTRUCTIONS FOR AUTHORS

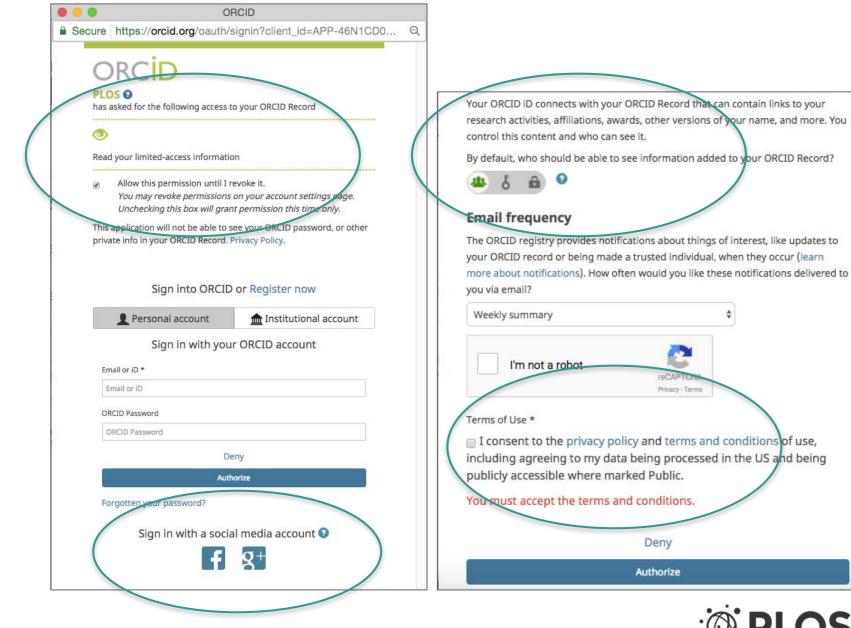


### Welcome to the PLOS ONE

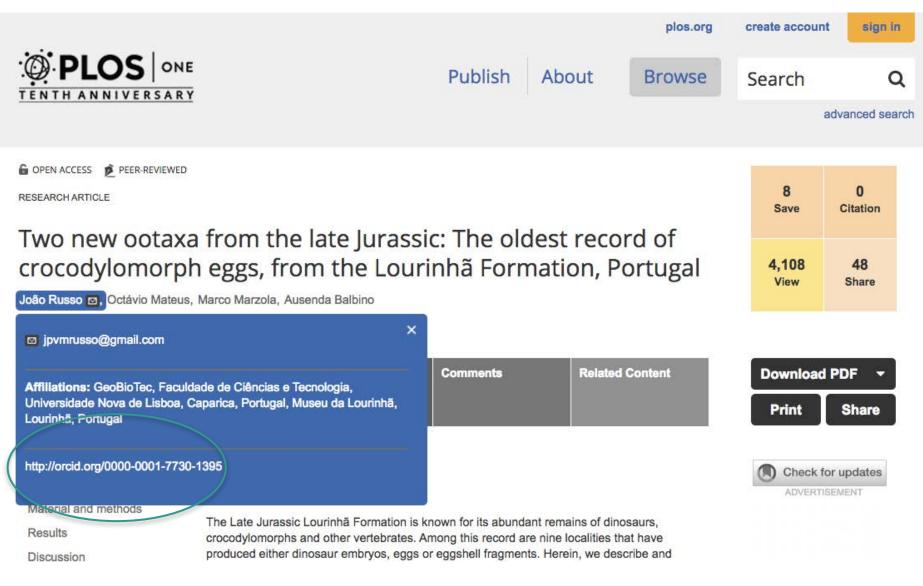
**Editorial Manager Submission System** 

Insert Special Character
Please Enter the Following
Username:
Password:
Author Login Reviewer Login Editor Login Publisher Login
Or Login via: (D) What is ORCID?
Send Login Details Register Now Login Help
PLOS requires an ORCID iD for all corresponding authors and encourages all
users to have one. To register for an ORCID iD and/or link your iD to your Editorial Manager account, click the green iD button. See this quick video
tutorial for more information.
Software Copyright © 2017 Aries Systems Corporation. Privacy Policy











## Publishers "Open Letter"

In January 2016, a coalition of publishers signed an Open Letter committing to start requiring ORCID IDs in 2016.

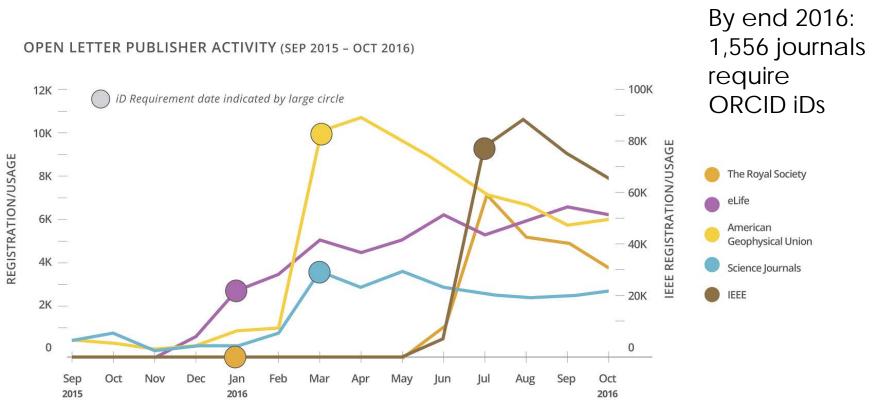
- 1. Implement best practices for ORCID collection
- 2. Commit to auto-update the ORCID records upon publication
- 3. Require ORCID IDs for corresponding authors and encourage for co-authors



## 8 original signatories, now 27!



## 27 Publishers requiring ORCID, and counting

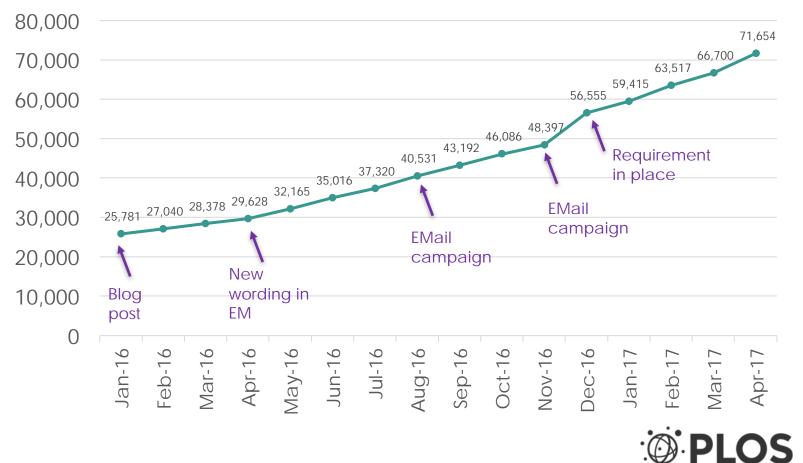


Since the open letter was published, over 250,000 articles have included ORCID iDs in their Crossref submission



## PLOS' sustained campaign

Authenticated ORCIDs





An open standard for expressing roles intrinsic to research









## CRediT: a taxonomy of contributions

Conceptualization

Methodology

Software

Validation

Formal Analysis

Investigation

Resources

Data Curation

Writing - Original Draft Preparation

Writing – Review & Editing

Visualization

Supervision

**Project Administration** 

Funding Acquisition

- Includes but is not limited to traditional author roles
- Not intended to define authorship
- Human- and machinereadable

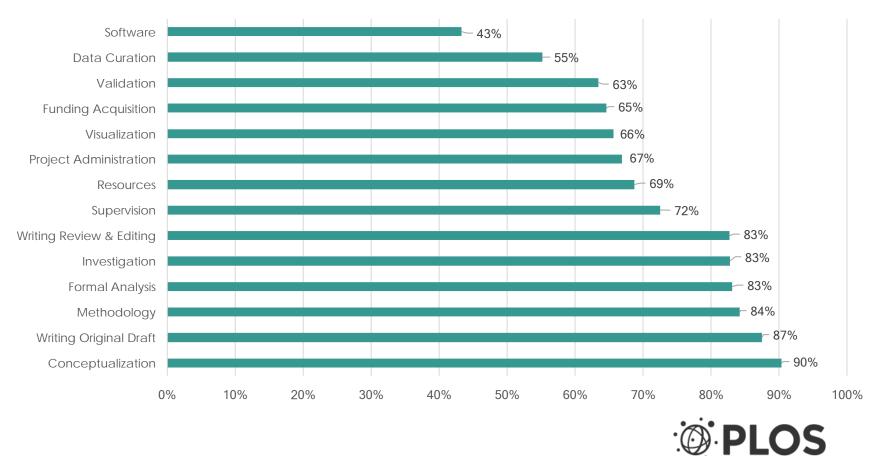
http://casrai.org/CRediT



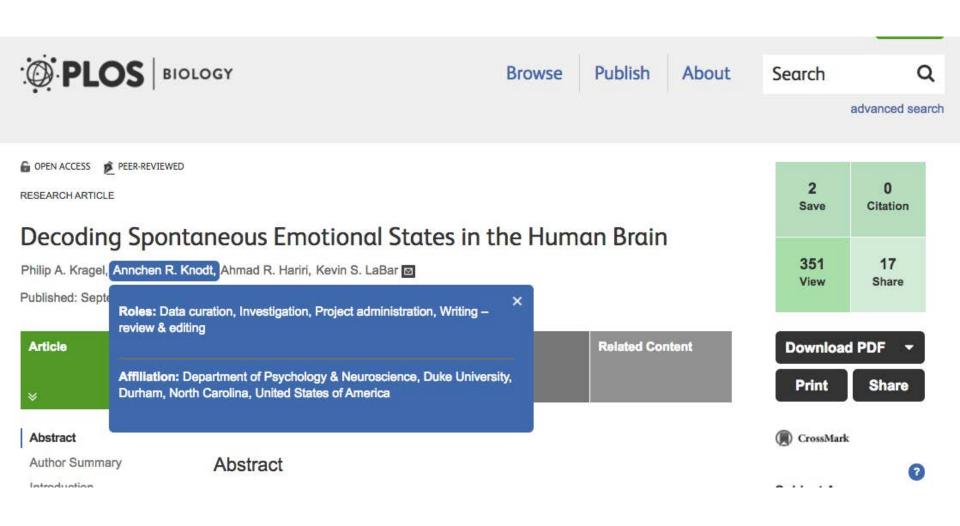


## Usage of CRediT taxonomy at PLOS

Frequency of use per contributor role



PLOS ONE submissions (n=3,833) – 92% at least one answer







- PLOS has been using CRediT since summer 2016 and requires ORCID for corresponding authors since Dec 2016.
  - All authors are encouraged to use ORCID

New Results	New	Results	
-------------	-----	---------	--

#### Transparency In Authors' Contributions And Responsibilities To Promote Integrity In Scientific Publication

Marcia McNutt, D Monica Bradford, D Jeffrey Drazen, R. Brooks Hanson, D Bob Howard,
 Kathleen Hall Jamieson, Veronique Kiermer, Michael Magoulias, Emilie Marcus,
 Barbara Kline Pope, Randy Schekman, Sowmya Swaminathan, Peter Stang, Inder Verma
 Inder Verma

This article is a preprint and has not been peer-reviewed [what does this mean?].

1.00	2000 C	20		182	22
• •	bs	**	<b>'</b> 2	e1	- I
-	20	-			-

Info/History Metrics

#### orr natory the

#### Abstract

In keeping with the growing movement in scientific publishing toward transparency in data and methods, we argue that the names of authors accompanying journal articles should provide insight into who is responsible for which contributions, a process should exist to confirm that the list is complete, clearly articulated standards should establish whether and when the contributions of an individual justify authorship credit, and those involved in the generation of scientific knowledge should follow these best practices. To accomplish these goals, we recommend that journals adopt common and transparent standards for authorship, outline responsibilities for corresponding authors, adopt the CRediT (Contributor Roles Taxonomy) methodology for attributing contributions, include this information in article metadata, and encourage authors to use the digital persistent identifier ORCID. Furthermore, we suggest that research institutions have regular open conversations on authorship criteria and ethics and that funding agencies adopt ORCID and accept CRediT. Scientific societies should further authorship transparency by promoting these recommendations through their meetings and publications programs.



#### **Subject Areas**

#### **All Articles**

Preview PDF

Animal Behavior and Cognition Biochemistry Bioengineering Bioinformatics Biophysics Cancer Biology Cell Biology Clinical Trials Developmental Biology Ecology Epidemiology Evolutionary Biology

## Data citations



FORCE11 » Groups » Joint Declaration of Data Citation Principles - FINAL

#### JOINT DECLARATION OF DATA CITATION PRINCIPLES - FINAL

When citing please use: Data Citation Synthesis Group: Joint Declaration of Data Citation Principles. Martone M. (ed.) San Diego CA: FORCE11; 2014 [/datacitation].

#### ENDORSEMENT LIST

#### PREAMBLE

Sound, reproducible scholarship rests upon a foundation of robust, accessible data. For this to be so in practice as well as theory, data must be accorded due importance in the practice of scholarship and in the enduring scholarly record. In other words, data should be considered legitimate, citable products of research. Data citation, like the citation of other evidence and sources, is good research practice and is part of the scholarly ecosystem supporting data reuse.







## Data Citation: credit for data producers and collectors

- Force11 Data Citation Principles
- Minimum Requirements
  - author names, repository name, date + persistent unique identifier (such as DOI or URI)
- citation should link to the dataset directly via the persistent identifier
- comprehensive, machine-readable landing pages for deposited data
- guidance to authors to include data in references





https://www.force11.org/group/joint-declaration-data-citation-principles-final



To ensure every researcher, at any phase of their career, or at any institution, will have seamless access to Persistent Identifiers (PIDs) for their research artefacts and their work will be uniquely attributed to them.



THOR Knowledge Hub @ project-thor.readme.io

# ORCID EMBL-EBI

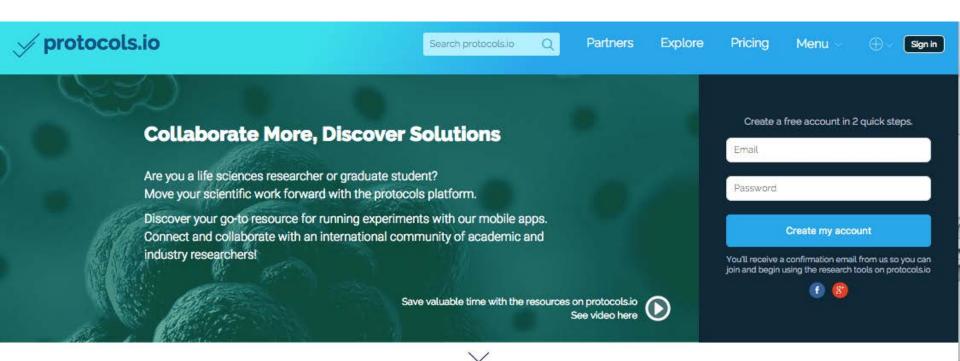


## THOR project

EU-funded project: Technical Infrastructures for Humans and Objects of Research

THOR's goal is to ensure that every researcher, at any phase of their career, or at any institution, will have seamless access to Persistent Identifiers (PIDs) for their research artefacts and their work will be uniquely attributed to them





#### When you sign up for protocols, you have the ability to:

#### For Researchers

- Log results with ease and save time
- Reproduce the experiment
- Correct and find solutions
- Edit and modify protocols dynamically with the ability to ask questions along the way
- Publish and collaborate privately or publicly
- Get credit for your scientific work

#### protocols.io

An up-to-date open access repository of science methods and a collaborative protocolcentered platform.

#### For Vendors & Publishers

- Bring your protocols to life on mobile & web for active work at the bench
- Improve your published methods based on the
- researcher-shared knowledge
- Help increase reproducibility of the published research



GOPEN ACCESS 💋 PEER-REVIEWED

RESEARCH ARTICLE

#### Commensal bacteria and essential amino acids control food choice behavior and reproduction

Ricardo Leitão-Gonçalves 📷, Zita Carvalho-Santos 📷, Ana Patrícia Francisco 📷, Gabriela Tondolo Fioreze, Margarida Anjos, Célia Baltazar, Ana Paula Elias, Pavel M. Itskov, Matthew D. W. Piper, Carlos Ribeiro o

Published: April 25, 2017 • https://doi.org/10.1371/journal.pbio.2000862

Article	Authors	Metrics	Comments	Related Content
*				

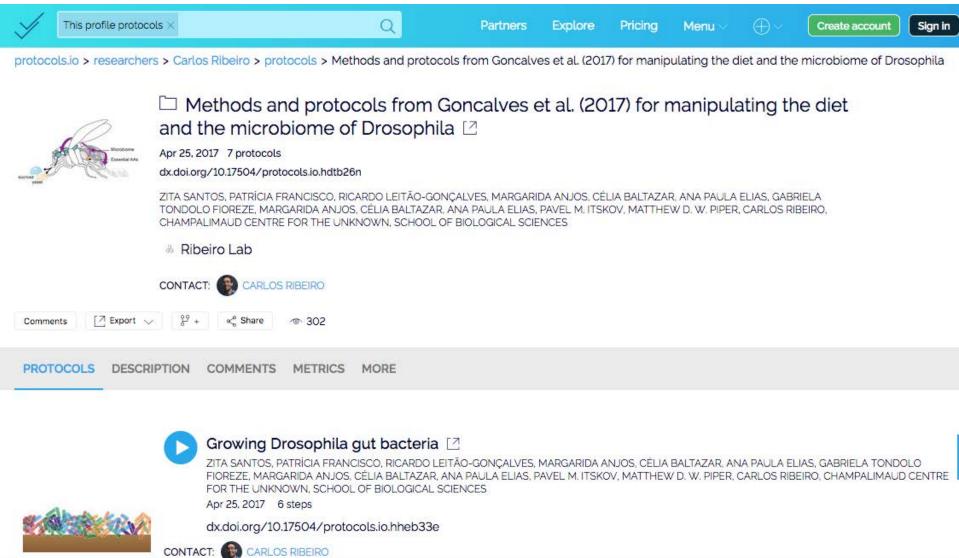
#### Abstract

Author summary	Abstract
Introduction	
Results	Choosing the right nutrients to consume is essential to health and wellbeing across species However, the factors that influence these decisions are poorly understood. This is particular
Discussion	1
Materials and methods	
Supporting information	Materials and methods
Acknowledgments	waterials and methods
References	Methods and protocols for Drosophila rearing, media preparare are available as a collection in protocols.io dx.doi.org/10.175
Reader Comments (0)	are available as a conection in protocols.to dx.doi.org/10.170
Media Coverage (11)	and reproductive output. This demonstrates now the interaction of specific nutrients with the
Figures	microbiome can shape behavioral decisions and life history traits.

#### aterials and methods

hods and protocols for Drosophila rearing, media preparations, and microbial manipulations available as a collection in protocols.io dx.doi.org/10.17504/protocols.io.hdtb26n.





search protocols, groups and more Q				Partners	Explore	Pri	icing Me	anu⊳ (	⊕~	Create ac	count Sign i
											save in journ
Lipids Amino acids Holidic media (HM) preparation Apr 25, 2017 12 steps ZiTA SANTOS, PATRICIA FRANCISCO, RICARDO LEITÃO-GONÇALVES, MARGARIDA ANJOS, CÉLIA B BALTAZAR, ANA PAULA ELIAS, PAVEL M. ITSKOV, MATTHEW D. W. PIPER, CARLOS RIBEIRO, CHAMI Holidic medium CONTACT: CARLOS RIBEIRO											)/12 STEPS MPLETED
cond part of the medium preparation											
6 temperature. Make s read more	Step 8										/×O
		pending on the nut ered water until the		nt of the me	edium add	all the	following co	omponents	s. At the (	end, if requi	red, add milliC
	Q	eferred to as	holidic medium	-all Aas (HUNTaa)	-eAAs (HUNTaa)		neAAs IUNTaa)	-single eA (HUNTaa)	AA (or	eAAs + 1xTyr r 2xTyr) IUNTaa)	1xPro
7 Schott flask with a read more tep Depending on the nutrient content of the medium add all the following components. At the end, if	305		holidic	(HUNTaa)		()		(HUNTaa)	NA (or ) (H		
7 Schott flask with a read more Rep Depending on the nutrient content of the medium add all the following components. At the end, if	Es No	eferred to as	holidic medium (HUNTaa)	(HUNTaa)	(HUNTaa)	( -	HUNTaa)	(HUNTaa)	NA (or ) (H	r 2xTyr) IUNTaa) I0.45 ml(a)	1xPro (HUNTaa)
<sup>29</sup> Depending on the nutrient content of the medium add all the following components. At the end, if	Es No aci	eferred to as ssential amino acids on-essential amino	holidic medium (HUNTaa) 60.51 ml	(HUNTaa) 0 ml	(HUNTaa) 0 ml	(H	HUNTaa) 40.45 ml(a)	(HUNTaa) 140.45 ml(a,c)	0 r	r 2xTyr) IUNTaa) I0.45 ml(a)	1xPro (HUNTaa) 140.45 ml(a) 0.27 g
7 Schott flask with a read more 8 Depending on the nutrient content of the medium add all the following components. At the end, if required, add milliQ fi read more	Es No aci	eferred to as ssential amino acids on-essential amino ids	holidic medium (HUNTaa) 60.51 ml 60.51 ml	(HUNTaa) 0 ml 0 ml	(HUNTaa) 0 ml 97.52 ml(a	(H	IUNTaa) 40.45 ml(a) ml	(HUNTaa) 140.45 ml(a,c) 0 ml	0 r	r 2xTyr) IUNTaa) I0.45 ml(a) ml 5.11 ml(a)	1xPro (HUNTaa) 140.45 ml(a) 0.27 g proline(f)
<ul> <li>Schott flask with a read more</li> <li>Depending on the nutrient content of the medium add all the following components. At the end, if required, add milliO fi read more</li> <li>ombining both solutions</li> <li>When the first part of the medium has cooled down enough (you should be able to touch the sides of</li> </ul>	Es No aci L-C Nu Vit	eferred to as ssential amino acids on-essential amino ids Glutamate ucleic acids & Lipids tamins	holidic međium (HUNTaa) 60.51 ml 60.51 ml 15.13 ml 8 ml 14 ml	(HUNTaa) 0 ml 0 ml 0 ml 8 ml 14 ml	(HUNTaa) 0 ml 97.52 ml(a 24.38 ml(a 8 ml 14 ml	(H	HUNTaa) 40.45 ml(a) ml 5.11 ml(a) ml 4 ml	(HUNTaa) 140.45 ml(a,c) 0 ml 35.11 ml( 8 ml 14 ml	(or (Hi 14 0 r (a) 35 8 r 14	r 2xTyr) IUNTaa) I0.45 ml(a) ml i.11 ml(a) ml I ml	1xPro (HUNTaa) 140.45 ml(a) 0.27 g proline(f) 35.11 ml(a) 8 ml 14 ml
<ul> <li>Schott flask with a read more</li> <li>Depending on the nutrient content of the medium add all the following components. At the end, if required, add milliQ fi read more</li> <li>Combining both solutions</li> <li>When the first part of the medium has cooled down enough (you should be able to touch the sides of</li> </ul>	Es No aci L-C Nu Vit	eferred to as ssential amino acids on-essential amino ids Glutamate ucleic acids & Lipids tamins blic acid	holidic medium (HUNTaa) 60.51 ml 60.51 ml 15.13 ml 8 ml 14 ml 1 ml	(HUNTaa) 0 ml 0 ml 0 ml 8 ml 14 ml 1 ml	(HUNTaa) 0 ml 97.52 ml(a 24.38 ml(a 8 ml 14 ml 1 ml	(H 14 i) 0 i) 35 8 14 14 1	HUNTaa) 40.45 ml(a) ml 5.11 ml(a) ml 4 ml ml	(HUNTaa) 140.45 ml(a,c) 0 ml 35.11 ml( 8 ml 14 ml 1 ml	(or ) (H 14 0 r (a) 35 8 r 14 17	r 2xTyr) (UNTaa) (0.45 ml(a) ml 5.11 ml(a) ml i ml ml ml	1xPro (HUNTaa) 140.45 ml(a) 0.27 g proline(f) 35.11 ml(a) 8 ml 14 ml 1 ml
<ul> <li>Schott flask with a read more</li> <li>Depending on the nutrient content of the medium add all the following components. At the end, if required, add milliQ fi read more</li> <li>Combining both solutions</li> <li>When the first part of the medium has cooled down enough (you should be able to touch the sides of</li> </ul>	Es No aci L-C Nu Vit	eferred to as ssential amino acids on-essential amino ids Glutamate ucleic acids & Lipids tamins blic acid cetic acid buffer	holidic medium (HUNTaa) 60.51 ml 60.51 ml 15.13 ml 8 ml 14 ml 1 ml 100 ml	(HUNTaa) 0 ml 0 ml 0 ml 8 ml 14 ml 1 ml 100 ml	(HUNTaa) 0 ml 97.52 ml(a 24.38 ml(a 8 ml 14 ml 1 ml 100 ml	(H 14 1) 0 35 8 14 1 1 10	IUNTaa) 40.45 ml(a) ml 5.11 ml(a) ml 4 ml ml 00 ml	(HUNTaa) 140.45 ml(a,c) 0 ml 35.11 ml( 8 ml 14 ml 1 ml 100 ml	(or )) (or 14: 0 r (a) 35: 8 r 14: 17: 10: 10: 10: 10: 10: 10: 10: 10	r 2xTyr) UUNTaa) 10.45 ml(a) ml 5,11 ml(a) ml I ml ml 00 ml	1xPro (HUNTaa) 140.45 ml(a) 0.27 g proline(f) 35.11 ml(a) 8 ml 14 ml 1 ml 100 ml
<ul> <li>7 Schott flask with a read more</li> <li>Schott flask with a read more</li> <li>Depending on the nutrient content of the medium add all the following components. At the end, if required, add milliO fi read more</li> <li>Combining both solutions</li> <li>When the first part of the medium has cooled down enough (you should be able to touch the sides of bottle, roughly 60 'C read more</li> </ul>	Es No aci L-C Nu Vit Fo Ac	eferred to as ssential amino acids on-essential amino ids Glutamate ucleic acids & Lipids tamins blic acid buffer eferred to as	holidic medium (HUNTaa) 60.51 ml 60.51 ml 15.13 ml 8 ml 14 ml 1 ml	(HUNTaa) 0 ml 0 ml 0 ml 8 ml 14 ml 1 ml 100 ml	(HUNTaa) 0 ml 97 52 ml(d 24 38 ml(d 8 ml 14 ml 1 ml 100 ml	(H 14 1) 0 35 8 14 1 1 10	HUNTaa) 40.45 ml(a) ml 5.11 ml(a) ml 4 ml ml	(HUNTaa) 140.45 ml(a,c) 0 ml 35.11 ml( 8 ml 14 ml 1 ml 100 ml	(a) 35 (a) 14 (a) 35 (a) 14 14 0 r 8 r 14 10 10	r 2xTyr) (UNTaa) (0.45 ml(a) ml 5.11 ml(a) ml i ml ml ml	1xPro (HUNTaa) 140.45 ml(a) 0.27 g proline(f) 35.11 ml(a) 8 ml 14 ml 1 ml 100 ml
<ul> <li>7 Schott flask with a read more</li> <li>Step Depending on the nutrient content of the medium add all the following components. At the end, if required, add milliQ fi read more</li> <li>Combining both solutions</li> <li>When the first part of the medium has cooled down enough (you should be able to touch the sides of bottle, roughly 60 'C read more</li> </ul>	Es. No aci L-C Nu Vit Fol Ac Es aci	eferred to as ssential amino acids on-essential amino ids Glutamate ucleic acids & Lipids tamins viic acid betic acid buffer eferred to as ssential amino ids	holidic medium (HUNTaa) 60.51 ml 60.51 ml 15.13 ml 8 ml 14 ml 1 ml 100 ml -single eAA	(HUNTaa) 0 ml 0 ml 0 ml 8 ml 14 ml 1 ml 100 ml (h) -folic a	(HUNTaa) 0 ml 97.52 ml(a 24.38 ml(a 8 ml 14 ml 1 ml 100 ml cid -me aa) (HU	(H 14 1) 0 1) 35 8 14 1 1 10 10 tals	IUNTaa) 40.45 ml(a) ml 5.11 ml(a) ml 4 ml ml 00 ml -nc.ac. & lij	(HUNTaa) 140.45 ml(a,c) 0 ml 35.11 ml( 8 ml 14 ml 1 ml 100 ml	(or)) (H (H) 0 r (a) 35 8 r 14 1 r 10 0 l 0 l	r 2xTyr) UNTaa) 10.45 ml(a) ml 5.11 ml(a) ml i ml ml 10 ml -vitamins	1xPro (HUNTaa) 140.45 ml(a) 0.27 g proline(f) 35.11 ml(a) 8 ml 14 ml 1 ml 100 ml
<ul> <li>Schott flask with a read more</li> <li>Depending on the nutrient content of the medium add all the following components. At the end, if required, add milliQ fi read more</li> <li>Combining both solutions</li> <li>When the first part of the medium has cooled down enough (you should be able to touch the sides of bottle, roughly 60 °C read more</li> </ul>	Es. No aci L-C Nu Vit Fol Ac Es aci	eferred to as ssential amino acids on-essential amino ids Glutamate ucleic acids & Lipids tamins blic acid betic acid buffer eferred to as ssential amino ids on-essential amino	holidic medium (HUNTaa) 60.51 ml 60.51 ml 15.13 ml 8 ml 14 ml 1 ml 100 ml -single eAA (HUNTaa)	(HUNTaa) 0 ml 0 ml 0 ml 8 ml 14 ml 1 ml 100 ml (h) -folic a (HUNT	(HUNTaa) 0 ml 97.52 ml(a 24.38 ml(a 8 ml 14 ml 100 ml cid -me aa) -me (HU nl 60.5	(H 14 1) 0 1) 35 8 14 1 1 10 tals NTaa)	IUNTaa) 40.45 ml(a) ml 5.11 ml(a) ml 4 ml 00 ml -nc.ac. & lij (HUNTaa)	(HUNTaa) 140.45 ml(a,c) 0 ml 35.11 ml( 8 ml 14 ml 1 ml 100 ml 100 ml	(a) 35 (a) 35 (a) 35 (a) 14 (a) 35 (a) 14 17 10 (a) 10 (a) 10 (a) 10 (a) 10 (b) 10 (c)	r 2xTyr) UNTaa) 10.45 ml(a) ml 5.11 ml(a) ml 6 ml ml 10 ml -vitamins (HUNTaa)	1xPro (HUNTaa) 140.45 ml(a) 0.27 g proline(f) 35.11 ml(a) 8 ml 14 ml 1 ml 100 ml holidic medium (FLYAA)



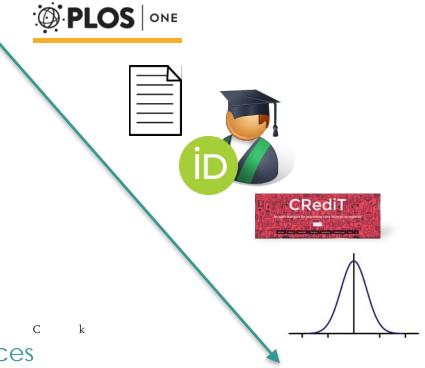
### Publishers' tools to facilitate better credit

- Citations distributions
- ✓ ORCID
- ✓ CRediT taxonomy
- ✓ Data citations
- ✓ Protocols
- ✓ Preprints

Raise awareness

Promote and facilitate better practices

Enable a machine-readable ecosystem





# Who's accountable?



# By the time an author submits to a journal it's too late...



## Data stewardship & sharing is spreading

- Other publishers are updating their data sharing policies and requiring a DAS
  - Nature, Science, Royal Society & Hindawi most recently
- Private funders have implemented policies requiring that data is made openly available.
  - Bill and Melinda Gates Foundation and Wellcome Trust (F1000 platforms)
  - Wellcome, HHMI, and NIH created the Open Science Prize to reward and make public the value of open, shared data.
- Government agencies have implemented or are exploring policies that facilitate data sharing.
  - Data Management plans as standard
  - National Institutes of Health (NIH), European Medical Association, European Commission and Research Council UK (RCUK)
- Academic institutions such as Lausanne, Cambridge University, University college London provide additional infrastructure and support for researchers to share data.
  - EU LEARN Project



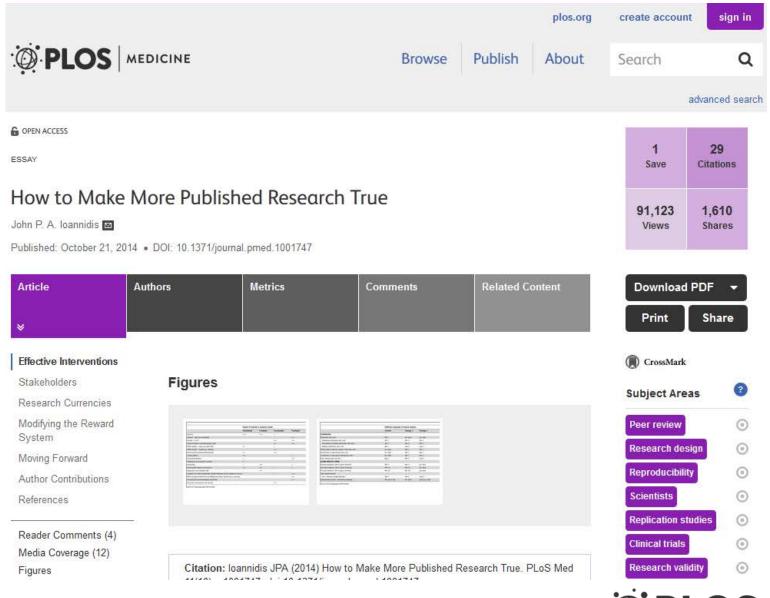
#### Solutions

- Open Access to articles and data
  - (that enables reuse CC BY, CC0)
- Separate the process of publication from evaluation
  - Make information openly available sconer (e.g. preprints)
  - PLOS-ONE style assessment (not in first, interest & novelty later)
    - Publish negative and confirmatory studies
- Open, signed Continuous peer review
  - More collective, community based review
- Incentivise openness, collaboration, reliability and sharing
  - Reward Reviewers
  - Reward open behaviour by researchers



• Reward all types of outputs – not just articles

- Apply the scientific method to scholarly communication itself
  - Meta-research research about the research process
  - Publically available data on metrics, indicators, evaluation
  - Independent scrutiny
- Align policies between funders, publishers, institutions
  - Data management as standard (& Data Access Committees)
  - Reduce the burden on researchers
  - Incentivise all players (sticks and carrots)
  - Monitor progress towards common goals
- Create global community standards for open science
  - Community standards for data & metadata sharing
  - NISO, FORCE11, COPE, TOP guidelines, Leiden Manifesto, HEFCE report on metrics, Reporting Standards
- Build the infrastructure to support open science
  - Interoperable publicly available platforms (EU Science Cloud)
  - New submission and reviewing tools that foster openness and collaboration, and do so earlier
  - The means to track and link all types of outputs
    - Persistent identifiers for researchers, funders, institution properties ORCID, FundRef, DOIs for data etc





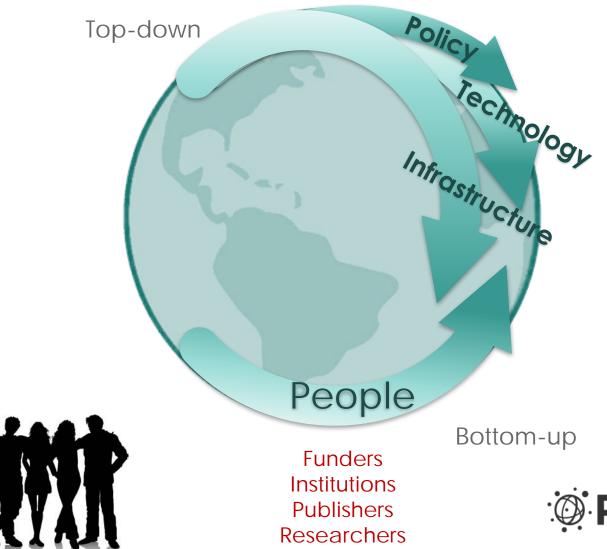
# Who's accountable?



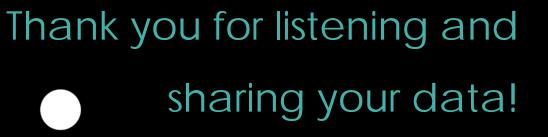
## ...we all are!



#### Cultural Change







cmaccallum@plos.org

orcid.org/0000-0001-9623-2225

Thanks to: PLOS: Meg Byrne Veronique Kiermer Emma Ganley Helen Atkins Patrick Polischuk

İD

CRediT: Amy Brand Liz Allen

"[Why this paper] was chosen for inclusion in our discussion is the fact that the actual data values in spreadsheet format is also available from the PLOS ONE website. You can download this and look at the data yourself... They used a Kruskal-Wallis test which is absolutely correct indeed."

STATISTICS COURSE INSTRUCTOR

