

A PRACTICAL GUIDE TO PRACTICING OPEN SCIENCE

Chris Gorgolewski
@ChrisFiloG

What is Open Science?

- Transparency
- Reproducibility
- Reusability
- Collaborations



WHO IS A PHD STUDENT?

WHO IS A POSTDOC?

WHO IS A TENURE TRACK
PROFESSOR?

Why should you care about open science

- Collaborations
- Grant funding requirements
- Job opportunities
- Community



Hey Krzysztof,

Hope you're well. I recently found your details online and I wanted to follow up with an email to let you know about some of the awesome [opportunities](#) we're hiring for in our deployed engineering function here in London (and some of our international offices too). Not sure if you remember but I actually dropped you a few emails back in may of last year. Although you were interested you mentioned that you weren't currently looking for a job in industry. Never the less I wanted to re-engage to see how things were going and if you might be open to exploring other [opportunities](#). From the info I've got on your profile it seems like your research marries up well with palantir and it's pretty evident from your [github](#) that your still and avid coder. How would you feel about taking your expertise into industry and own problems across a broad area of environments?

There's a high possibility you've been contacted by a few other [recruiting](#) agencies so rather than bombarding with you with more words I've put a few links below for you to check out that will really give you insight into the company. I totally understand you may not be looking for a new position. The best people never are :-). But if you'd be open to having a chat, totally informal of course, I'd really appreciate your time.

Apply here online. Applications should include a CV, research plan, and cover letter, which may be up to 4 pages. We encourage including the following in the cover letter:

1. Provide links to open source code (e.g. through GitHub or other code repository) that you are most proud of

Roadmap

1. How to deal with data
2. How to deal with code
3. How to deal with papers

**OPEN
SCIENCE**

**D
A
T
A**

**C
O
D
E**

**P
A
P
E
R
S**

DATA

Consent forms

- General Data Protection Regulation
 - 25 May 2018



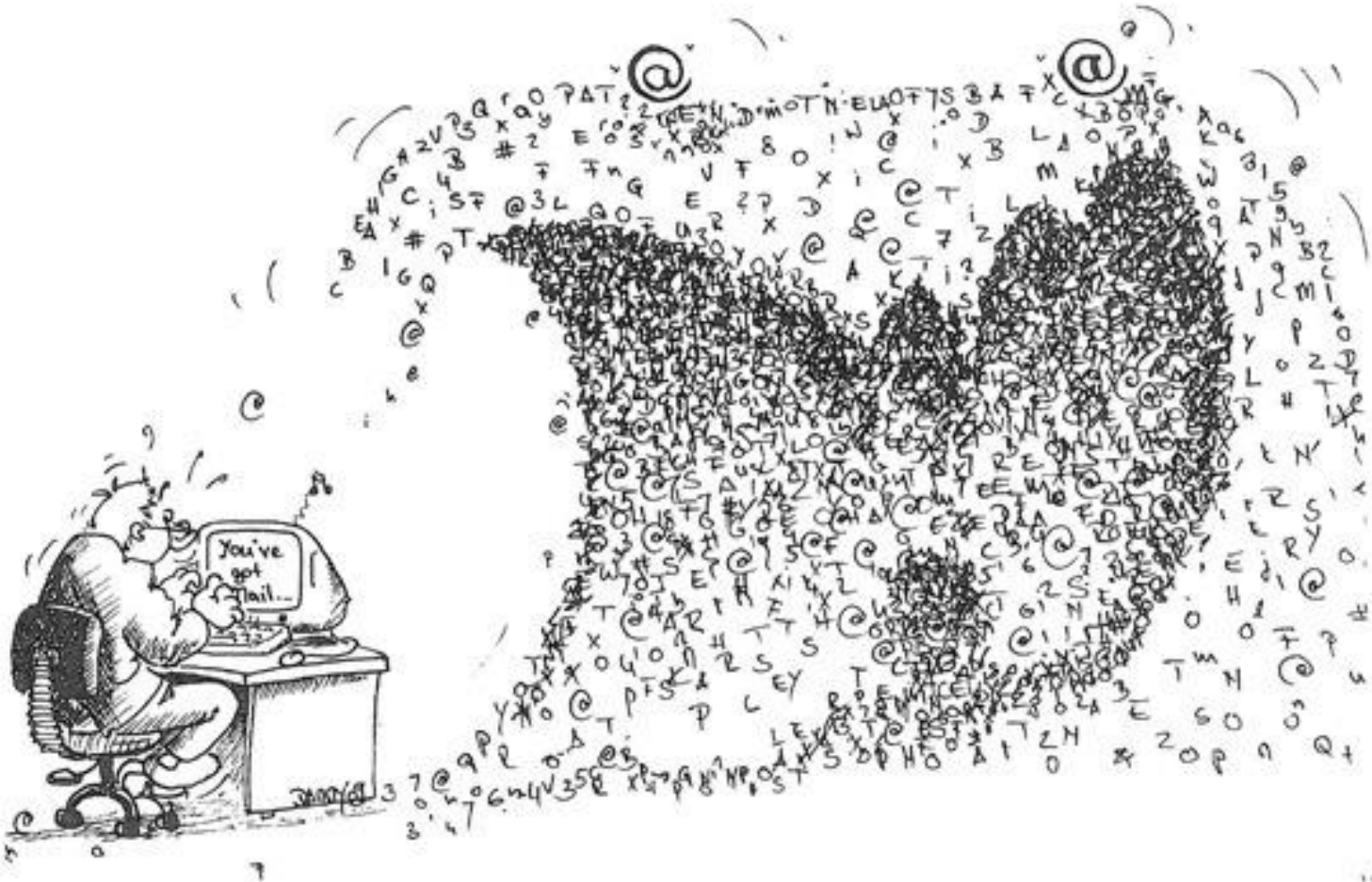
- Open Brain Consent
 - <http://open-brain-consent.readthedocs.io>
 - Now in German!

Open Brain Consent

However, by using additional data linked to your name (for example brain scans obtained from your medical records) **one could potentially use your imaging or other information in our database back to you.**

If you change your mind and withdraw your consent to participate in this study (you can call <PI name> at <phone number> to do this), we will not collect any additional data about you. We will delete your data if you withdraw before it was deposited in the database. **However, any data and research results already shared with other investigators or the general public cannot be destroyed, withdrawn or recalled.**

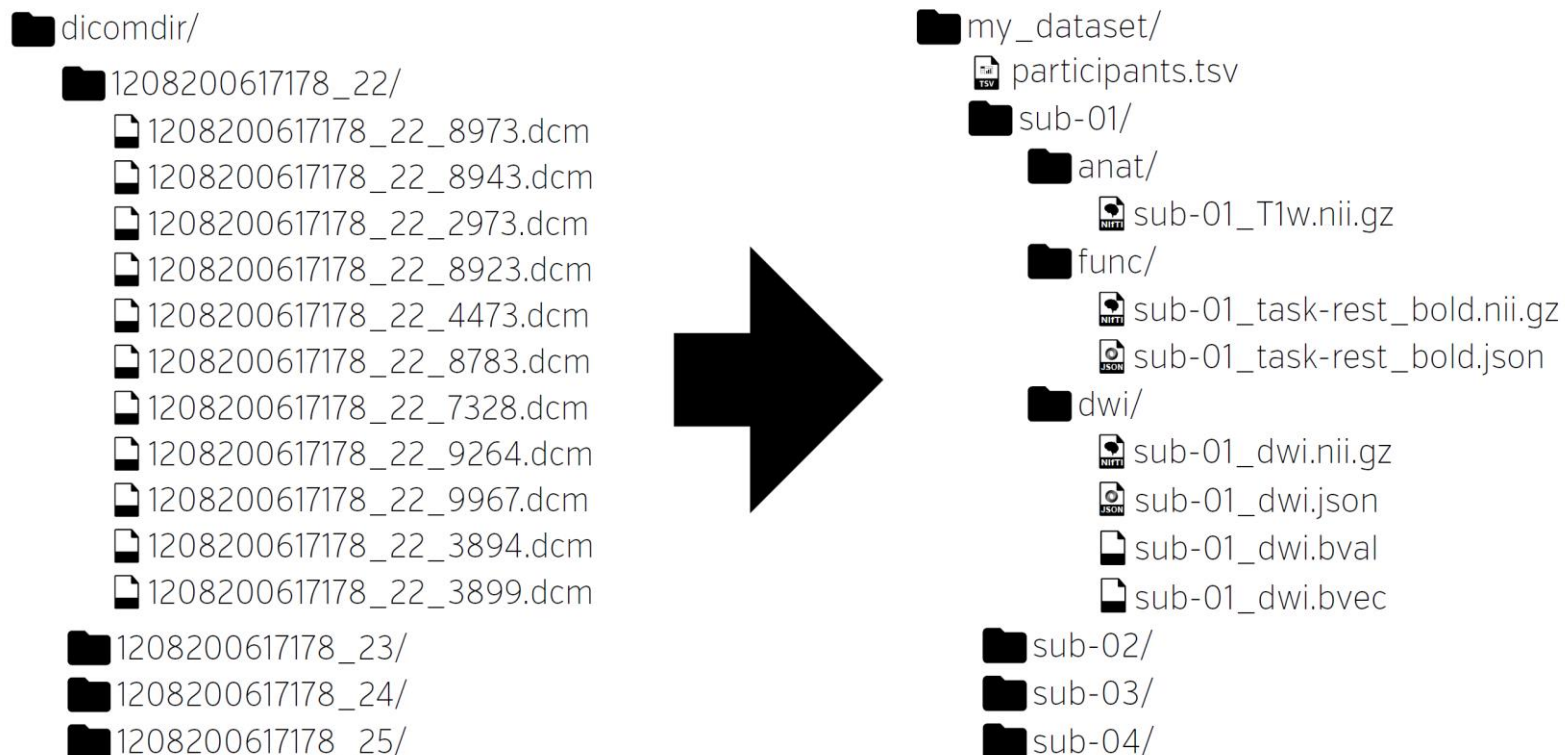
Data organization



Data organization

Data sharing with yourself from the future is
also a form of data sharing.
A very important one.

Brain Imaging Data Structure



BIDS.NEUROIMAGING.IO

Data repositories

- Curated:

- OpenfMRI.org




- FCP-INDI

- NITRC



INDI
International Neuroimaging
Data-Sharing Initiative



 The source for neuroinformatics tools & resources
 Neuroimaging data repository
 Cloud computing environment

- Uncurated:

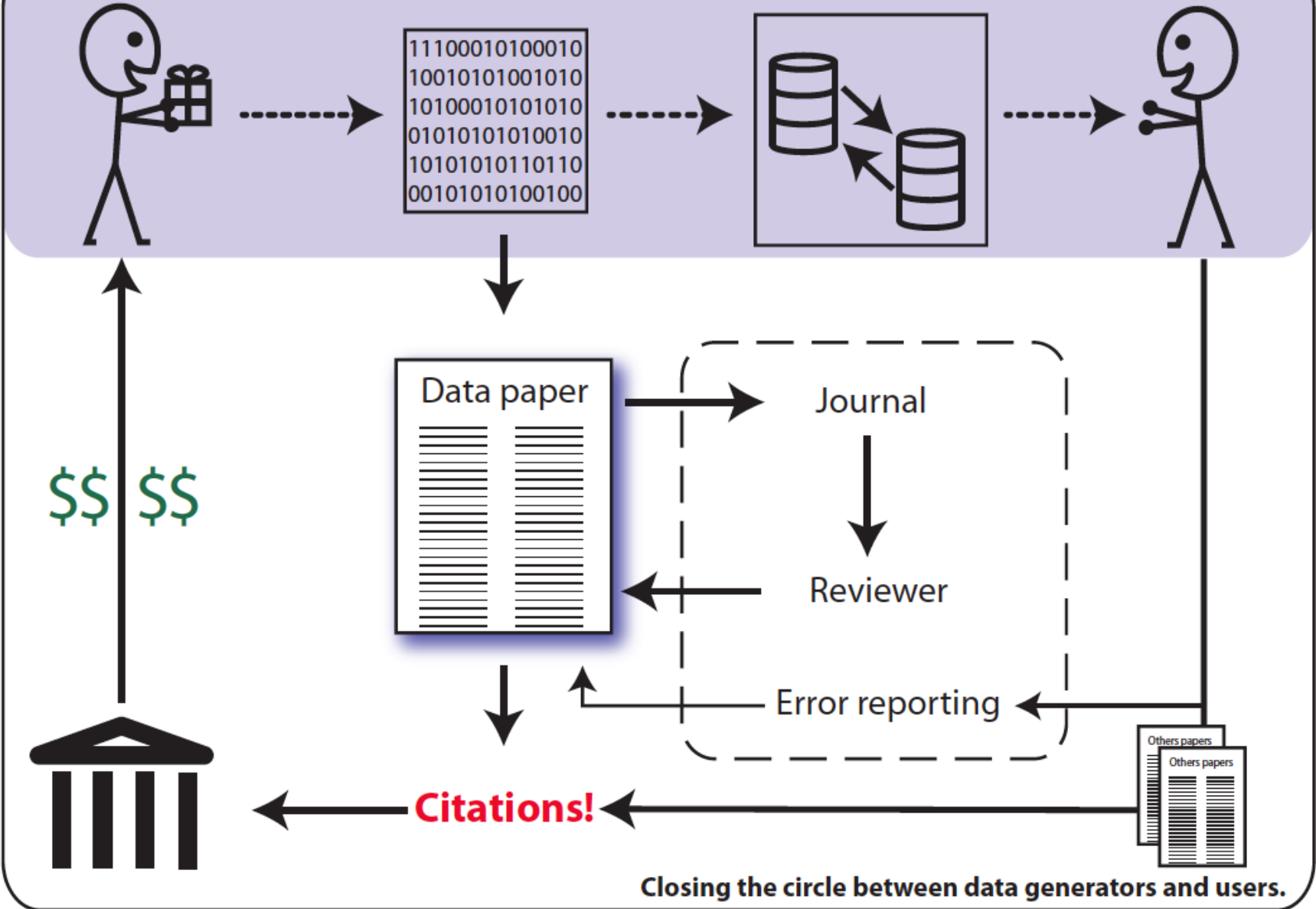
- Data Dryad

- Figshare

- Harvard Dataverse

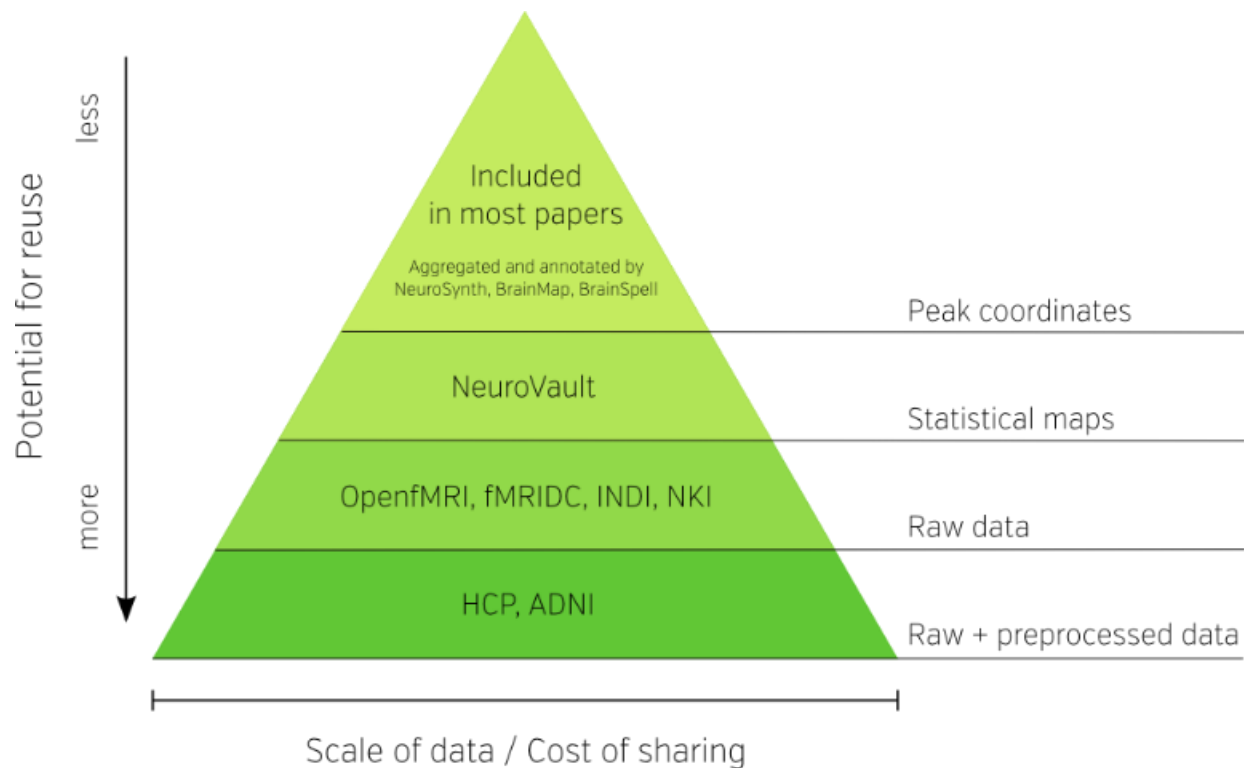


Current model of open data sharing.



Data sharing fears

- Scooping and Errors



CODE

Where to put your code?

- GitHub
 - For day to day updates



- Zenodo
 - For DOI, coauthorship and long-term preservation



May 21, 2017

Software Open Access

Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in Python. 0.13.1

[ID](#) Gorgolewski, Krzysztof J.; [ID](#) Esteban, Oscar; [ID](#) Ellis, David Gage; [ID](#) Notter, Michael Philipp; [ID](#) Ziegler, Erik; [ID](#) Johnson, Hans; [ID](#) Hamalainen, Carlo; Yvernault, Benjamin; Burns, Christopher; [ID](#) Manhães-Savio, Alexandre; [ID](#) Jarecka, Dorota; [ID](#) Markiewicz, Christopher J.; [ID](#) Salo, Taylor; [ID](#) Clark, Daniel; Waskom, Michael; Wong, Jason; Modat, Marc; [ID](#) Dewey, Blake E; [ID](#) Clark, Michael G.; [ID](#) Dayan, Michael; Loney, Fred; Madison, Cindee; [ID](#) Gramfort, Alexandre; [ID](#) Keshavan, Anisha; Berleant, Shoshana; Pinsard, Basile; Goncalves, Mathias; [ID](#) Clark, Dav; [ID](#) Cipollini, Ben; [ID](#) Varoquaux, Gael; [ID](#) Wassermann, Demian; [ID](#) Rokem, Ariel; [ID](#) Halchenko, Yaroslav O.; Forbes, Jessica; Moloney, Brendan; [ID](#) Malone, Ian B.; [ID](#) Hanke, Michael; Mordom, David; Buchanan, Colin; [ID](#) Pauli, Wolfgang M.; [ID](#) Huntenburg, Julia M.; [ID](#) Horea, Christian; Schwartz, Yannick; Tungaraza, Rosalia; [ID](#) Iqbal, Shariq; Kleesiek, Jens; Sikka, Sharad; Frohlich, Caroline; Kent, James; Perez-Guevara, Martin; Watanabe, Aimi; Welch, David; Cumba, Chad; Ginsburg, Daniel; [ID](#) Eshaghi, Arman; [ID](#) Kastman, Erik; Bougacha, Salma; Blair, Ross; [ID](#) Acland, Benjamin; [ID](#) Gillman, Ashley; [ID](#) Schaefer, Alexander; [ID](#) Nichols, B. Nolan; Giavasis, Steven; Erickson, Drew; Correa, Carlos; Ghayoor, Ali; Küttner, René; Haselgrove, Christian; Zhou, Dale; [ID](#) Craddock, R. Cameron; Haehn, Daniel; Lampe, Leonie; Millman, Jarrod; Lai, Jeff; Renfro, Mandy; Liu, Siqi; [ID](#) Stadler, Jörg; [ID](#) Glatard, Tristan; [ID](#) Kahn, Ari E.; [ID](#) Kong, Xiang-Zhen; [ID](#) Triplett, William; Park, Anne; McDermottroe, Conor; Hallquist, Michael; Poldrack, Russell; Perkins, L. Nathan; Noel, Maxime; [ID](#) Gerhard, Stephan; Salvatore, John; Mertz, Fred; [ID](#) Broderick, William; Inati, Souheil; Hinds, Oliver; Brett, Matthew; [ID](#) Durnez, Joke; Tambini, Arielle; Rothmei, Simon; [ID](#) Andberg, Sami Kristian; [ID](#) Cooper, Gavin; Marina, Ana; Mattfeld, Aaron; Urchs, Sebastian; Sharp, Paul; Matsubara, K; [ID](#) Geisler, Daniel; Cheung, Brian; [ID](#) Floren, Andrew; Nickson, Thomas; [ID](#) Pannetier, Nicolas; Weinstein, Alejandro; Dubois, Mathieu; Arias, Jaime; Tarbert, Claire; Schlamp, Kai; [ID](#) Jordan, Kesshi; Liem, Franz; Saase, Victor; Harms, Robbert; Khanuja, Ranjeet; Podranski, Kornelius; Flandin, Guillaume; [ID](#) Papadopoulos Orfanos, Dimitri; Schwabacher, Isaac; [ID](#) McNamee, Daniel; Falkiewicz, Marcel; [ID](#) Pellman, John; [ID](#) Linkersdörfer, Janosch; Varada, Jan; [ID](#) Pérez-García, Fernando; Davison, Andrew; Shachnev, Dmitry; [ID](#) Ghosh, Satrajit

0.13.1 (May 20, 2017)

- FIX: Make release compatible with conda-forge build process (<https://github.com/nipy/nipype/pull/2017>)
- ENH: Update some minimum versions in compliance with Debian Jessie (<https://github.com/nipy/nipype/pull/2017>)



Tweeted by 8

[See more details](#)

Available in

Publication date:

May 21, 2017

DOI:DOI: [10.5281/zenodo.581704](https://doi.org/10.5281/zenodo.581704)**Keyword(s):**[neuroimaging](#) [workflow](#) [pipeline](#)**Related identifiers:**

Supplement to:

<https://github.com/nipy/nipype/tree/0.13.1>**License (for files):**[Apache Software License 2.0](#)

How to test your code

- Testing
 - Smoke tests + continuous integration
 - Assertion tests



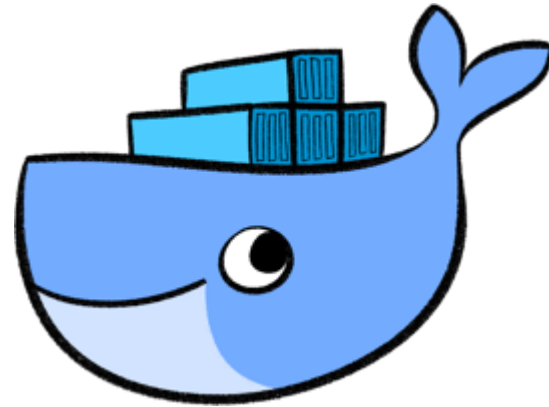
Reproducibility

reproducibility
=
shared code +
shared environment +
shared data

Sharing environments

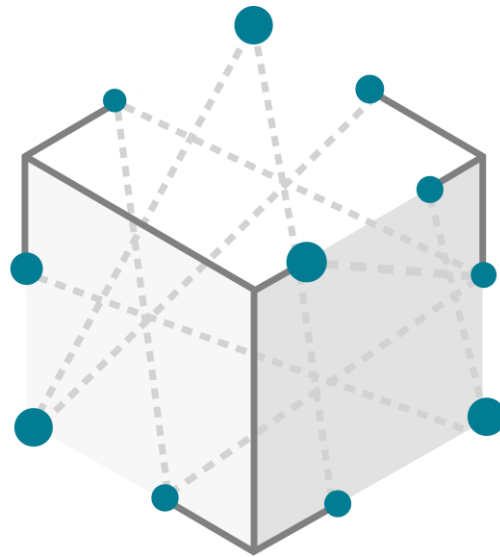
- All binary dependencies
- Environment variables
- Libraries
- etc.

Docker



<https://neurohackweek.github.io/docker-for-scientists/>

Reproducibility of neuroimaging analyses



OpenNEURO

a free online platform for sharing and analysis of neuroimaging data

Upload Dataset



1: Select

2: Rename

3: Issues

Your dataset is not a valid BIDS dataset. Fix the 1 Error and select your folder again.

VIEW 1 ERROR IN 1 FILE ^

Error: 1

VIEW 1 FILE

The compulsory file /dataset_description.json is missing. See Section 8.1 of the BIDS specification.

VIEW 3 WARNINGS IN 53 FILES v

SELECT FOLDER

Downl

Clic

Upload Dataset



1: Select

2: Rename

3: Issues

4: Disclaimer

5: Progress

ds000009_R2.0.1

34/856 files complete

uploading

sub-29_task-stopsignal_run-01_bold.nii.gz...

sub-29_dwi.nii.gz...

sub-29_T1w.json

Misconception 1: Quality of code

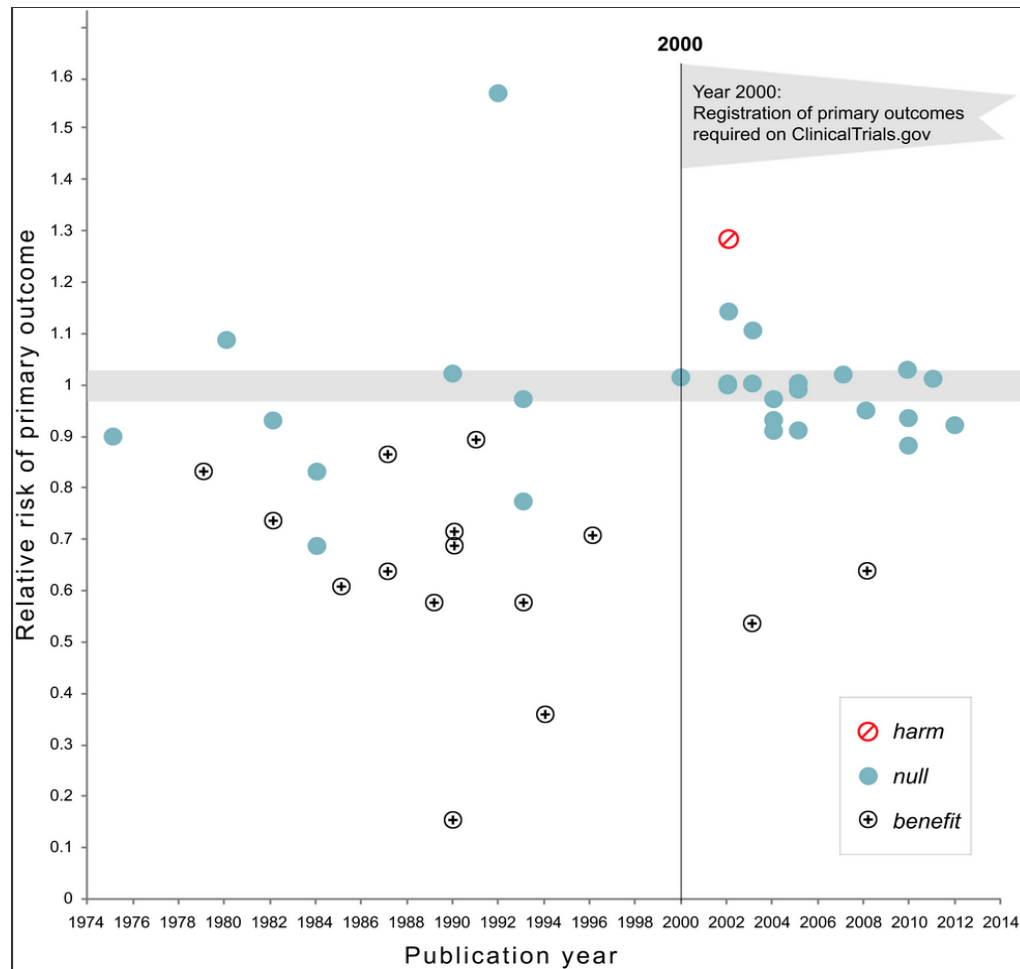
Lack of documentation and portability should not stop you from sharing your code.

Misconception 2: User support

1. You are not obliged to provide user support for your code if you share it.
2. Direct users to GitHub Issues or NeuroStars to avoid direct messages.

PAPERS

Preregistrations in clinical trials



Preregistrations

- OSF

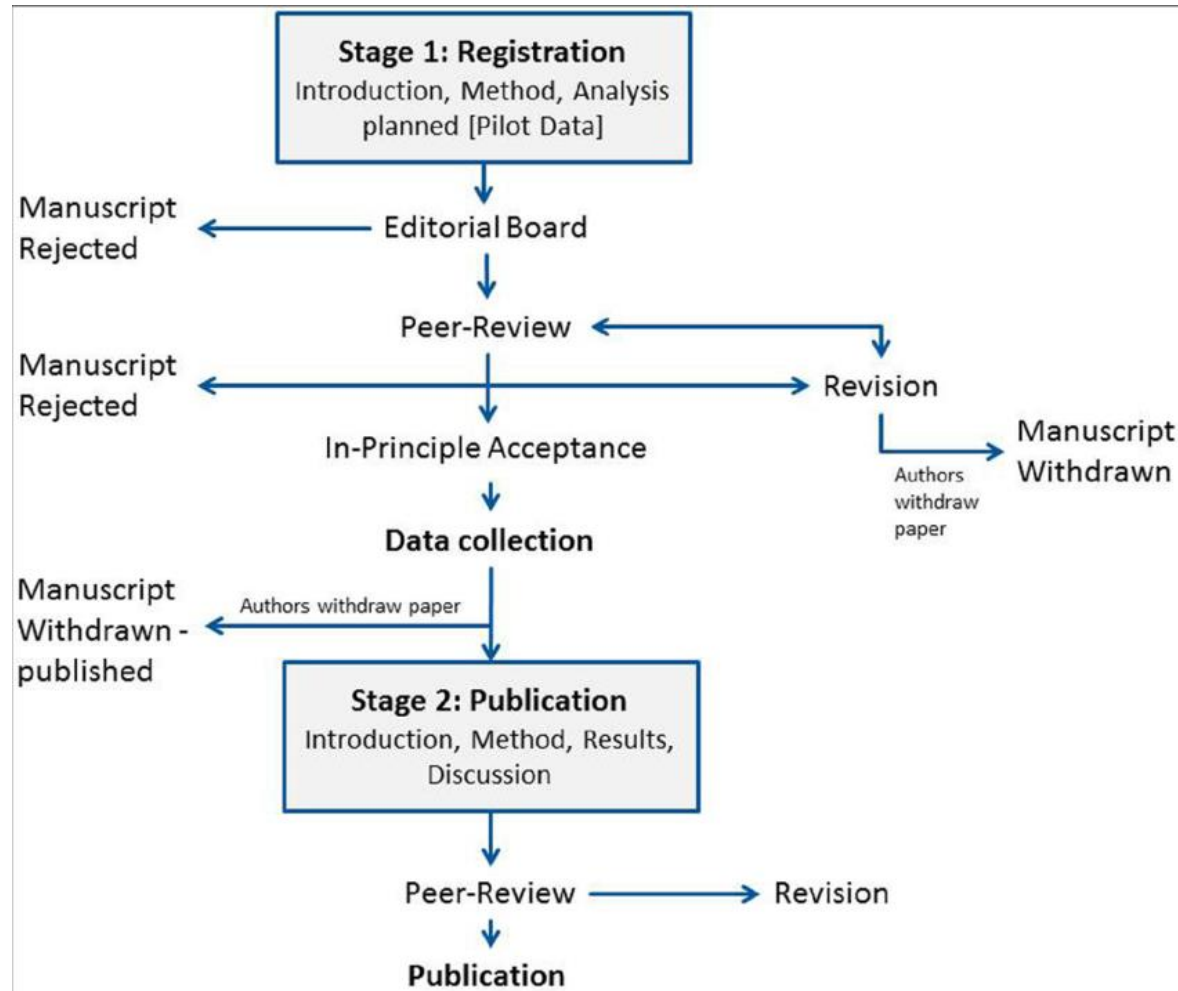


- AsPredicted



It's up to you how specific you want to be.

Registered reports



Preprints

- Way around non-open access journals
- A way to get feedback from the community
 - And use it to get your paper published
- Ability to share preliminary ideas

Preprints

- What makes a good preprint server?
 - Exposure
 - Comments
 - DOIs
 - Versions

bioRxiv
beta

THE PREPRINT SERVER FOR BIOLOGY

Preprint misconceptions

The journal will not accept a paper that already appeared as a preprint.

95% wrong

<http://www.sherpa.ac.uk/romeoinfo.html>

Reviews

- Academic Karma allows you to publish your reviews
 - Signed or not
 - Negotiating an agreement with authors when a review is solicited by the editor



Before accepting the review invitation, the reviewer would like to ask the authors if they are willing to agree to the following: (Set a field to blank if don't wish to specify.)

Post a preprint of their manuscript as submitted to journal

Yes

Post a revised preprint after addressing reviewers comments

Yes

Make source code openly available

Yes

Make data openly available

Yes

Comments

The reviewer is also willing to commit to: (Set a field to blank if don't wish to specify.)

Complete review within following period:

2 weeks

Agree to reviewing a revised preprint, if required

Yes

Publish review on Publons after manuscript is published

Yes

FINALLY...

The world is not black and white

It's many shades of grey

The world is not black and white



Not this kind of grey...

The world is not black and white

You don't need to go all in – pick
one thing
and start practicing it

Questions?

Demo

The screenshot displays the OpenNeuro website interface. At the top, there is a navigation bar with links for 'MY DASHBOARD', 'PUBLIC DASHBOARD', 'SUPPORT', 'FAQ', and 'UPLOAD DATASET'. The main content area is titled 'Deciphering consciousness' and includes the following information:

- Files:** 131, **Size:** 2.2GB, **Subjects:** 16, **Session:** 1
- Available Tasks:** balloon analog risk task
- Available Modalities:** T1w, inplaneT2, bold
- AUTHORS:** Schonberg TS, Fox CR, Mumford JA, Congdon E, Trepel C, Poldrack RA, Ronald McDonald
- README:** This dataset was obtained from the OpenfMRI project (<http://www.openfmri.org>).

On the right side, there is a 'BIDS Validation' section showing a green checkmark and the word 'Valid', along with a '3 WARNINGS' indicator. Below this is a 'Dataset File Tree' showing a hierarchical structure of files and folders, including 'dataset_description.json', 'participants.tsv', 'README', 'task-balloonanalogrisktask_bold.json', and subfolders 'sub-01' through 'sub-06'. Each file in the tree has options for 'UPDATE', 'DELETE', 'DOWNLOAD', and 'VIEW'.

At the bottom of the page, the text 'OpenNeuro.org - Poster #1677' is visible.