No country for old data: Increasing FAIR-ness of research outcomes through standardization and community-driven development

CSPD 2020

Peer Herholz (he/him)
Postdoctoral researcher - NeuroDataScience-ORIGAMI lab at MNI, McGill
- BRAMS
Member - BIDS, ReproNim, Brainhack, UNIQUE



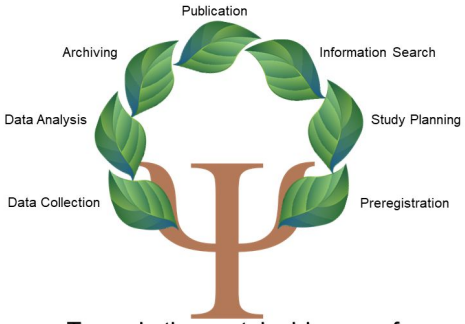
@peerherholz

December 07, 2020

A few things before we start ...

- to the folks with hearing impairments:
  - There will be live captions at the bottom of the screen. If something doesn't work, please let me/us know in the chat.
- to the folks with vision impairments:
  - All graphics (except those taken from publications) should be high in contrast and color-blind friendly. If there are problems, please let me/us know in the chat.
- to everyone:
  - I will try my best to speak loud and clearly. If you have troubles understanding me, please say so.
  - If you feel uncomfortable doing this/asking questions via the public chat, please send me a direct message.
- to everyone (continued):
  - I sincerely hope you feel comfortable indicating your pronouns, thus please add them to your name if you want to.
  - During the discussion/questions, please say your name as I (and others) would like to avoid mispronouncing it.
  - I try my best to present an objective overview. However, there's inherent bias to such presentations based on personal opinions and experience. Thus, I'm looking forward to openly discuss every point/comment with you all.
  - Let's all give our best to create an open, supportive and welcoming atmosphere for everyone.
- special thanks for pioneering work go to:
  - Kirstie Whitaker, Fernando Perez, Elizabeth DuPre, Gael Varoquaux

A few things before we start ...



Towards the sustainable use of psychological research data

Who here has reused data?

If so, what kind?

A data sharing nightmare in three parts … or “A data sharing carole”

https://youtu.be/N2zK3sAtr-4

It gets better! …. Does it really?

the classic

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request. **Code availability**

DATA AVAILABILITY STATEMENT

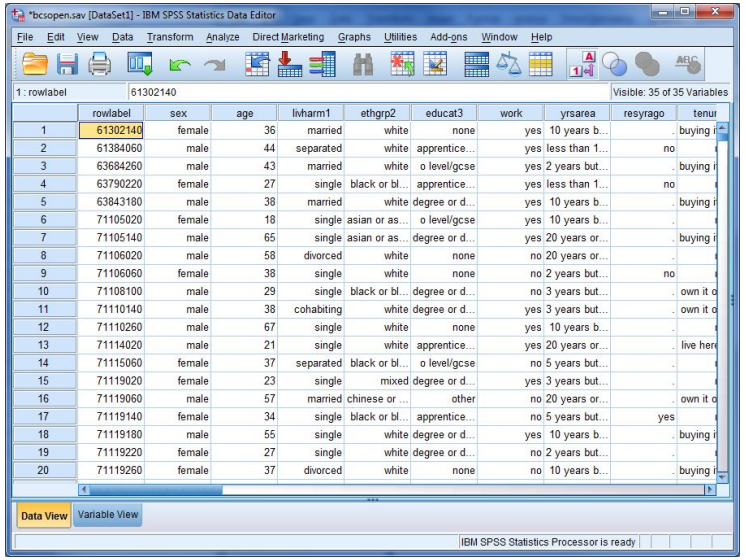
Code used to generate the analyses are available from the corresponding author upon reasonable request.

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Data Availability Statement Go to:

The datasets generated for this study are available on request to the corresponding author.

the solution?



https://bizlib247.files.wordpress.com/2014/04/data-editor-labels.png

"upon (reasonable) request" =
you might get it if I (a) can find it and (b) feel like it

if "data" is shared = here you go, have fun

The great baking analogy...

data as shared by author

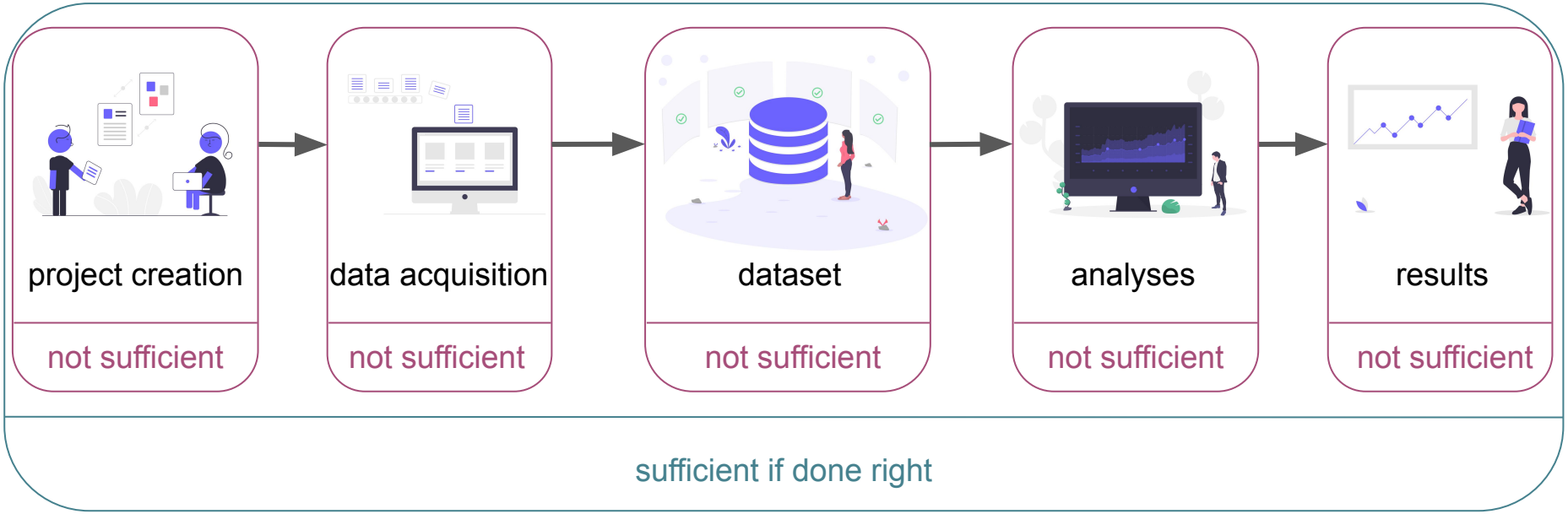
folks interested in the data

data/results in fancy journal

?

?

The great baking analogy...every ingredient AND the recipe are required



Please make sure to check https://undraw.co/

*F*indable *I*nteroperable *A*ccessible *R*eusable research outcomes

Findable
- data and metadata easily findable for both humans and machines (machine readable metadata *essential* for automatic discovery)

Accessible
- data can be easily obtained by humans as well as by machines, through well-defined and ideally standardized protocols
- conditions must be clearly established (license, reuse rights, etc.).
- when the data is inaccessible, the metadata should remain so in a way that the dataset remains discoverable

Interoperable
- possible to integrate different data into a single interface
- use of a common programming language, the use of controlled vocabulary, as well as non-proprietary formats
- can also be applied to metadata as well as references to other (meta)data

Re-Usable
- data and metadata following best practices -> optimized for reuse
- data and metadata must be richly detailed, as well as user licenses
- any information specific to the field of interest of the data in question must be made available as well.

https://ogsl.ca/wp-content/uploads/Fair-rectangle-en.png

- https://www.force11.org/fairprinciples
- https://www.go-fair.org/fair-principles/
- https://www.ands.org.au/working-with-data/fairdata
- https://www.nature.com/articles/sdata201618.pdf

Findable ***I***nteroperable ***A***ccessible ***R***eusable research outcomes

- FAIR is *fuzzy per design*
- a *general set of principles* that should *guide* research *across domains*
- FAIR has *natural variation across domains* (e.g. metadata in Psychology or Neuroscience)
- each domain *needs* to develop its own FAIR *standards and terms*, as no solution can fit all

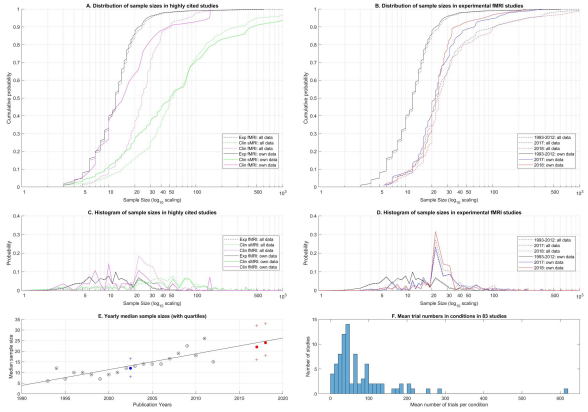
https://ogsl.ca/wp-content/uploads/Fair-rectangle-en.png

- https://www.force11.org/fairprinciples
- https://www.go-fair.org/fair-principles/
- https://www.ands.org.au/working-with-data/fairdata
- https://www.nature.com/articles/sdata201618.pdf

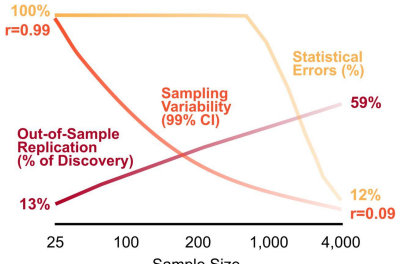
The four horsemen of … unFAIR research outcomes

Source: Public domain

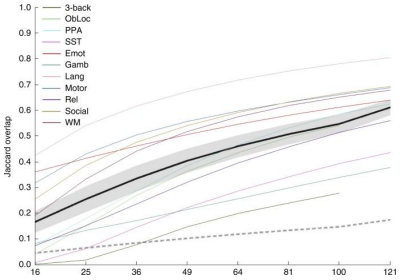
We need more data and we need it now, but FAIR



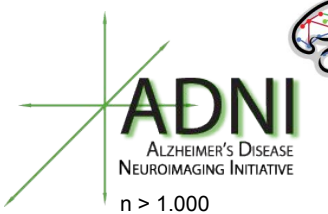
Szucs & Ioannidis (2020)



Marek et al. (2020)



Turner et al. (2018)



n > 1.000

n = 12.000

CONNECTOME
COORDINATION FACILITY

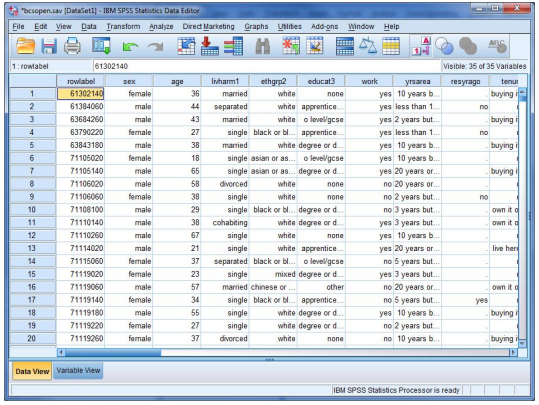
n ~ 10.000

biobank uk

Category ID	Description	Items
1014	Brain MRI	922
1005	Cognitive function summary	5
1004	Diet and alcohol summary	321
1002	Early life	13
1007	Education and employment	16
1017	Genomics	78
100113	Geographical and location	13
1015	Heart MRI	39
1019	Linked health outcomes	75
1016	Main abdominal MRI fields likely to be of interest to researchers.	22
1018	Mental health	186
1006	Physical measure summary	66
1001	Primary demographics	8
1003	Self-reported medical conditions	118

n = 40.000

FAIR data through standardization … *with Marie Kondo*

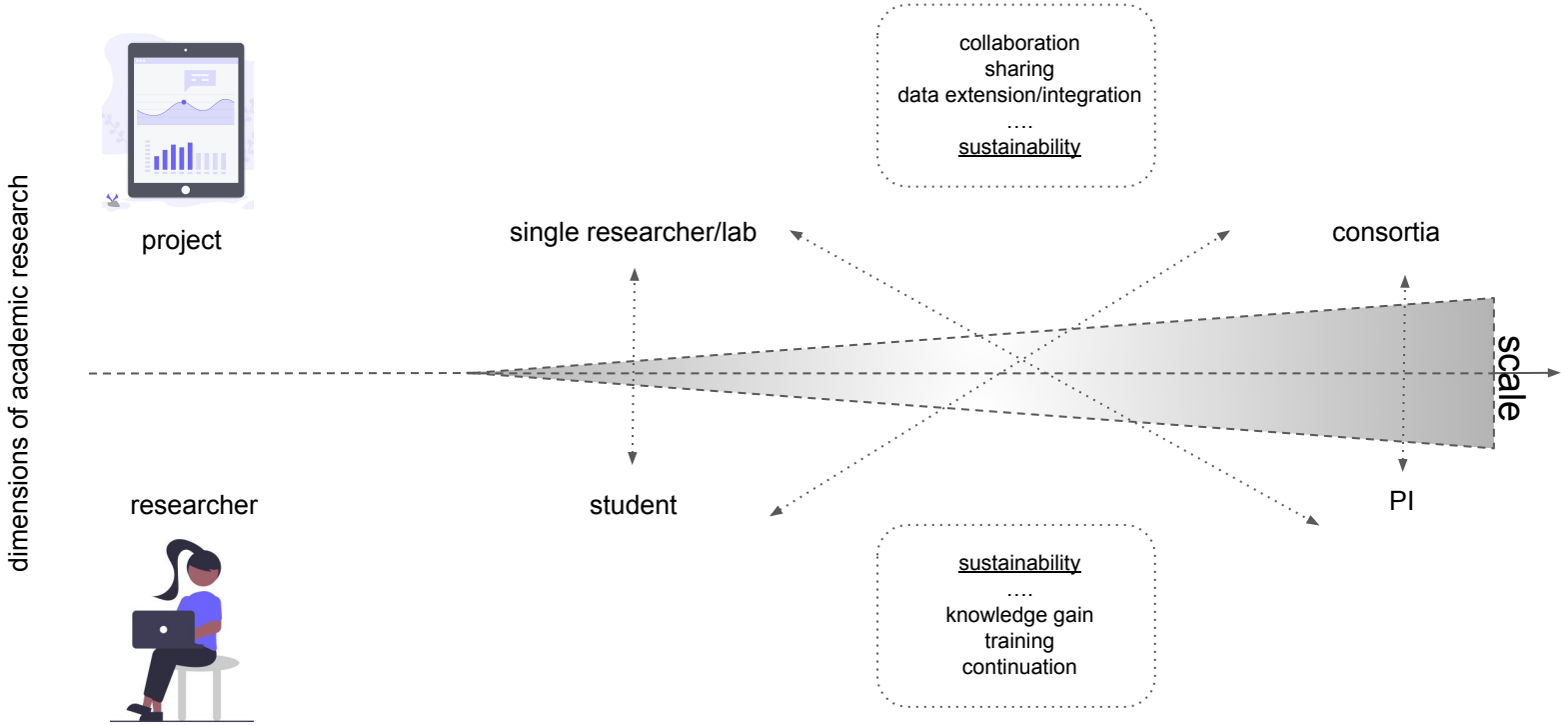


https://bizlib247.files.wordpress.com/2014/04/data-editor-labels.png

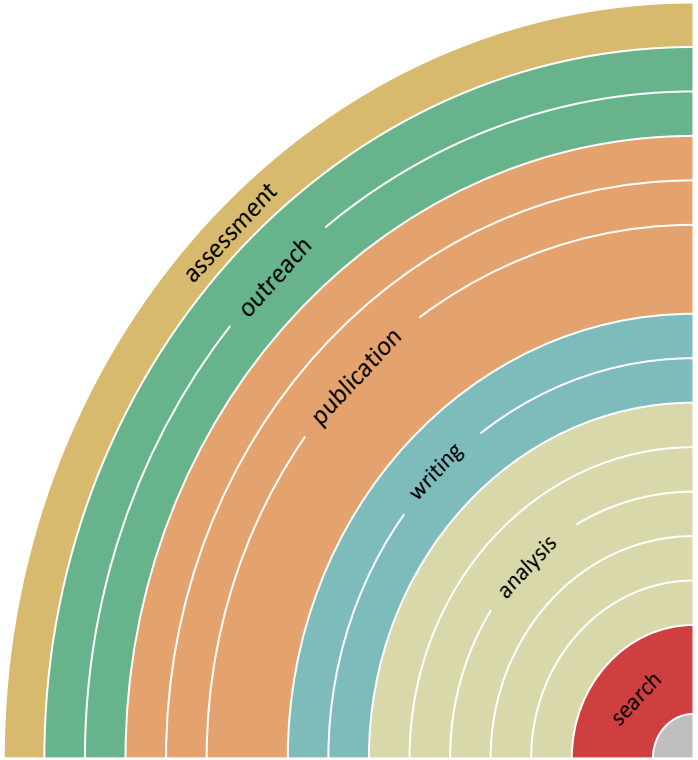
```
sub-control01/
    anat/
        sub-control01_T1w.nii.gz
        sub-control01_T1w.json
        sub-control01_T2w.nii.gz
        sub-control01_T2w.json
    func/
        sub-control01_task-nback_bold.nii.gz
        sub-control01_task-nback_bold.json
        sub-control01_task-nback_events.tsv
        sub-control01_task-nback_physio.tsv.gz
        sub-control01_task-nback_physio.json
        sub-control01_task-nback_sbref.nii.gz
    dwi/
        sub-control01_dwi.nii.gz
        sub-control01_dwi.bval
        sub-control01_dwi.bvec
    fmap/
        sub-control01_phasediff.nii.gz
        sub-control01_phasediff.json
        sub-control01_magnitude1.nii.gz
        sub-control01_scans.tsv
code/
    deface.py
derivatives/
README
participants.tsv
dataset_description.json
CHANGES
```

```
{
  "test": {
    "LongName": "Education level",
    "Description": "Education level, self-rated by participant",
    "Levels": {
      "1": "Finished primary school",
      "2": "Finished secondary school",
      "3": "Student at university",
      "4": "Has degree from university"
    }
  },
  "bmi": {
    "LongName": "Body mass index",
    "Units": "kg/m^2",
    "TermURL": "http://purl.bioontology.org/ontology/SNOMEDCT/60621009"
  }
}
```

FAIR for all? YES!



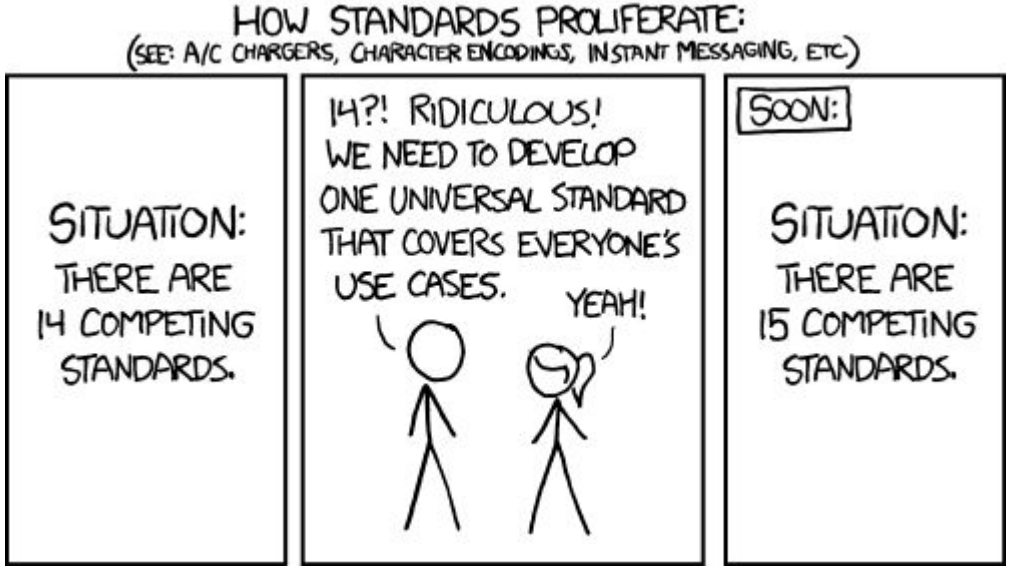
FAIR by default, from the beginning to the end and beyond



- adding alternative evaluation, e.g. with altmetrics
- communicating through social media, e.g. Twitter
- sharing posters & presentations, e.g. at FigShare
- using open licenses, e.g. CC0 or CC-BY
- publishing open access, 'green' or 'gold'
- using open peer review, e.g. at journals or PubPeer
- sharing preprints, e.g. at OSF, arXiv or bioRxiv
- using actionable formats, e.g. with Jupyter or CoCalc
- open XML-drafting, e.g. at Overleaf or Authorea
- sharing protocols & workfl., e.g. at Protocols.io
- sharing notebooks, e.g. at OpenNotebookScience
- sharing code, e.g. at GitHub with GNU/MIT license
- sharing data, e.g. at Dryad, Zenodo or Dataverse
- pre-registering, e.g. at OSF or AsPredicted
- commenting openly, e.g. with Hypothes.is
- using shared reference libraries, e.g. with Zotero
- sharing (grant) proposals, e.g. at RIO

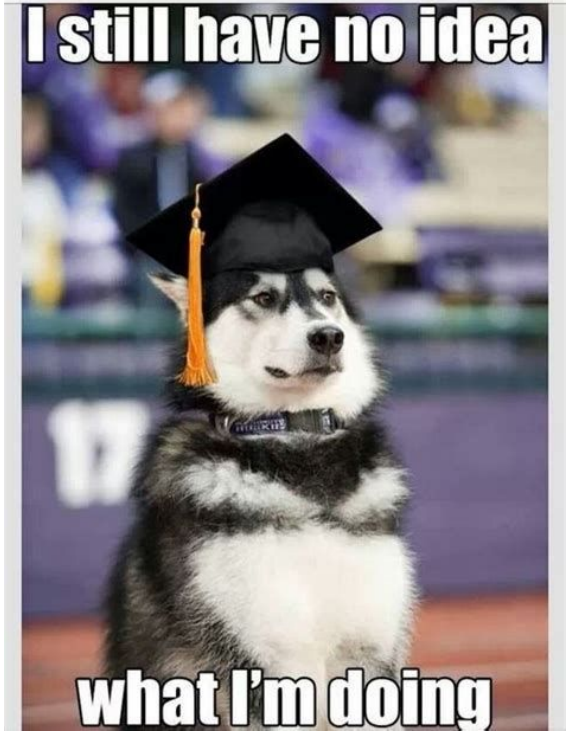
DOI: 10.5281/zenodo.1147025

FAIR data through standardization - problem I



https://imgs.xkcd.com/comics/standards.png

FAIR data through standardization - problem II



"The broad adoption of computers in experimental science has both enabled and obscured research" (Marwick, 2015)

- digital literacy, project/data management, FAIR, open science, etc. not part of most curricula (independent of field of study)
- vast majority across levels not trained in practical skills needed for sufficient and FAIR research

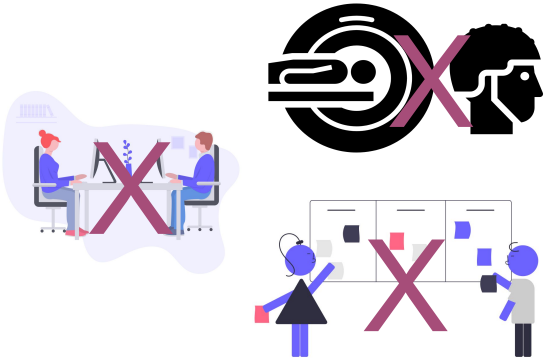
FAIR data through standardization - problem III

* researchers (looking at you students) are frustrated because their programme did not provide them with an realistic point of view regarding research workflows
→ the folks who do the analyses

* the people at the top (looking at you PIs) often expect comprehensive and exhaustive skills without providing the means and opportunities to learn them
→ the folks who don't do the analyses

FAIR data through standardization - problem III

the year 2020...



* there's a ton of open data out there with more participants than folks could have acquired themselves, there's open source software for everything, there are open and free cloud/remote computing possibilities, there are open and free lecture/tutorials/workshops/etc.
  → no one trained to use these opportunities

* PIs/institutes/departments "force" researchers to go the office, acquire data, etc.

* researchers (looking at you students) are frustrated because they can't acquire data, can't go to the lab to run analyses, can't discuss work with peers
  → a lot of projects/theses are effectively on hold since March

* PIs/institutes/departments not prepared for any of this
  → management problems get worse, major problems re grant applications, etc.

FAIR data through standardization - problem III (spoilers)

Offene Wissenschaft in der Zeit von Covid-19 – Ein Blaupause für die psychologische Forschung?

Julia Beitner[1], Garvin Brod[2,3], Benjamin Gagl[1], Dominik Kraft[4], und Martin Schultze[5]

[1] Allgemeine Psychologie I, Goethe-Universität Frankfurt am Main. Theodor-W.-Adorno-Platz 6, D-60323 Frankfurt am Main.

[2]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Rostocker Str. 6, D-60323 Frankfurt am Main.

[3]Pädagogische Psychologie, Goethe-Universität Frankfurt am Main. Theodor-W.-Adorno-Platz 6, D-60323 Frankfurt am Main.

[4]Neurokognitive Psychologie, Goethe-Universität Frankfurt am Main. Theodor-W.-Adorno-Platz 6, D-60323 Frankfurt am Main.

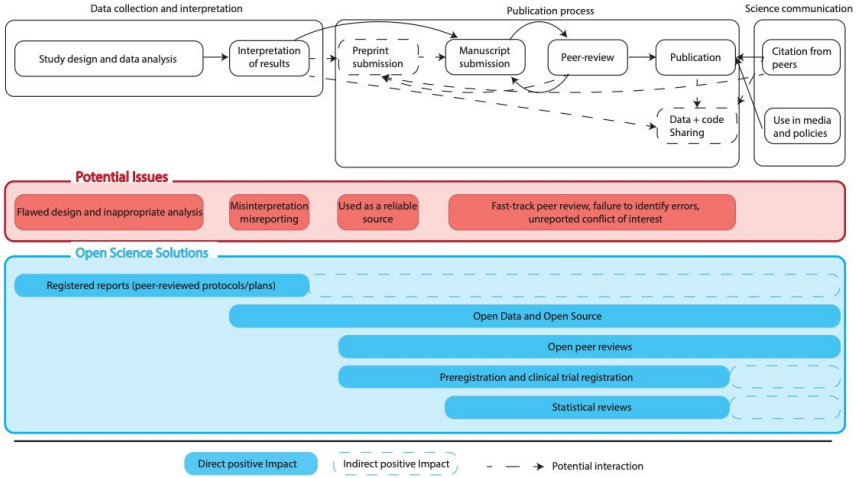
[5]Psychologische Methoden mit interdisziplinärer Ausrichtung, Goethe-Universität Frankfurt am Main. Theodor-W.-Adorno-Platz 6, D-60323 Frankfurt am Main.

Alle Open Science Initiative Frankfurt am Main. Autorenreihenfolge alphabethisch.

https://doi.org/10.31234/osf.io/sh8xg

Open Science Saves Lives: Lessons from the COVID-19 Pandemic

Lonni Besançon [1,2*], Nathan Peiffer-Smadja [3,4], Corentin Segalas [5], Haiting Jiang[6], Paola Masuzzo [7], Cooper Smout [7], Eric Billy[8], Maxime Deforet [9+] and Clémence Leyrat [5,10+]



https://doi.org/10.1101/2020.08.13.249847

FAIR data through standardization - problem I/II/III - the goal

“Das Internet ist Neuland.”

(“The internet is uncharted territory.”)

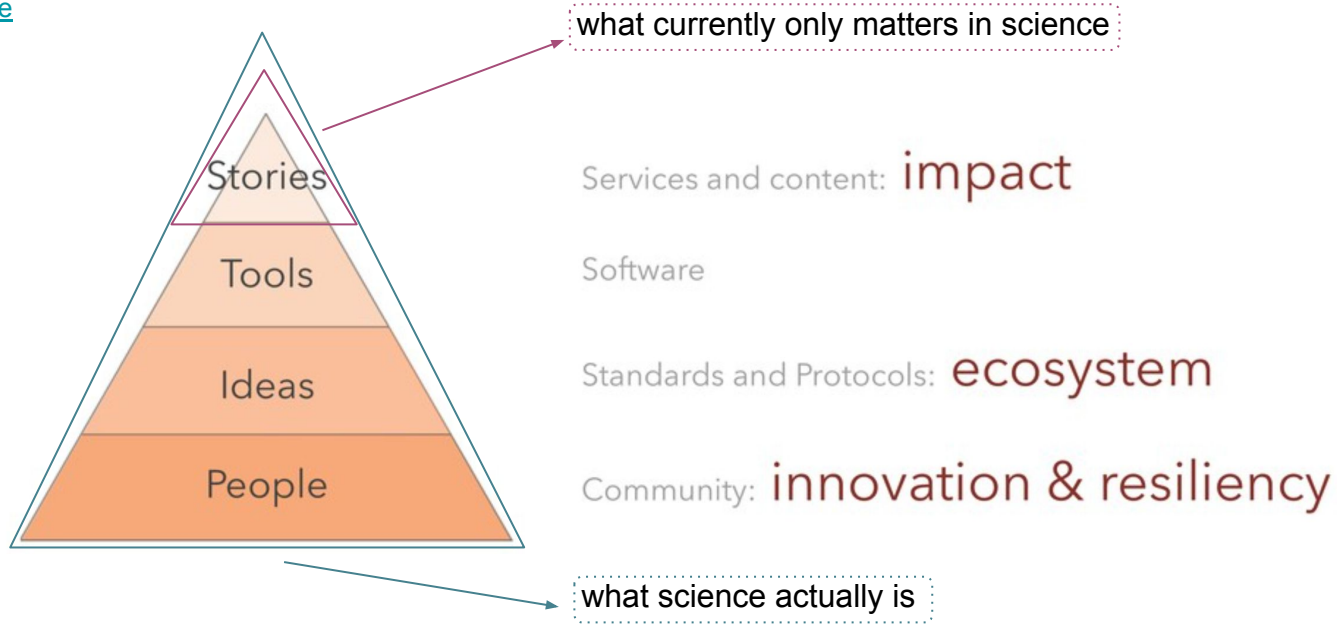
- How do I manage projects and data?
- Why do I have to conduct the x-th underpowered study with limited resources?
- Why can’t I access data and run analyses from everywhere?
- Why do I have to learn everything on my own and am dependent on the good will of my supervisors?

“The future is now.”

- I have a solid project and data management plan.
- I can use open data from thousands of participants.
- No matter where I am, I can work with my data.
- There is a large international, open and supportive community that provides help no matter what.

FAIR data through standardization - problem IV

*adapted from Fernando Perez and Elizabeth DuPre



"An article about computational science in a scientific publication is not the scholarship itself, it is merely advertising of the scholarship. The actual scholarship is the complete software development environment and the complete set of instructions which generated the figures."

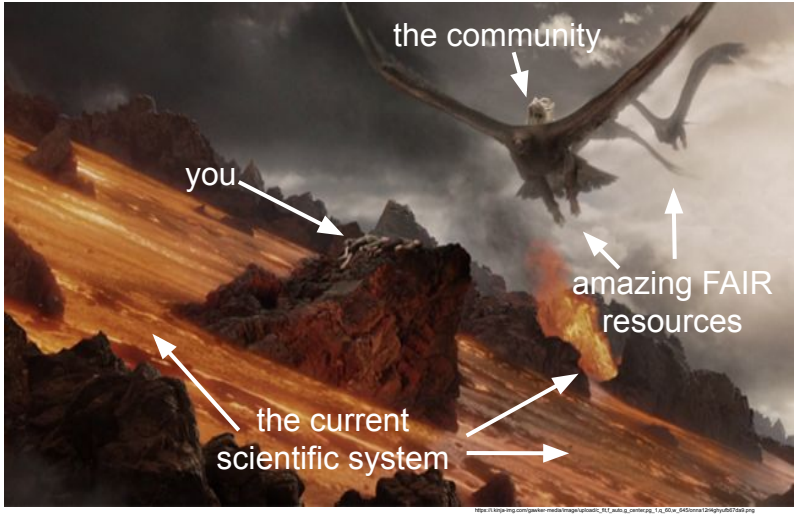
Buckhelt and Donoho
(paraphrasing John Claerbout)
WaveLab and Reproducible Research, 1995

FAIR data through standardization - problem IV

	open communities	classic science/academia
Credit	Distributed	PI & hierarchy
Output/artifacts	Continuous & Project-specific	Discrete papers
Collaborators	Fluid: professionals, volunteers, …	Structured, funding-dependent
Governance/ decision making	Open, community based	Top-down, PI
Authorship	Fluid, roles can evolve, no clear "first/ senior" author	Need to say more?
Peer review	Continuous, open, pervasive, friendly	The opposite
Value metric	Utility, need, impact	"Novel and transformative"

*adapted from Fernando Perez

FAIR data through standardization - problem IV



VS

community based governance

* The governance and decision making process is clearly structured, open to everyone from the community, actively seeking input from as many folks from as many backgrounds as possible, as well as allows to incorporate changes and updates in a continuous manner.

hierarchical top-down power structure

* A few powerful chosen ones make decisions that affect a tremendous amount of those “below” them, even though they are not aware of latest developments and haven’t worked with data since ages.

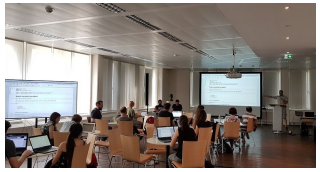
Rhetorical question: did I chose an analogy depicting old white men on purpose?

FAIR data through community-driven development of standards and beyond

*adapted from Elizabeth DuPre

- **openly** develop standards, materials, resources
- **consensus** in decision-making
- **empower** and **equip** community members
- communities **grow from conscious investment** rather than the availability of resources
- communities and science in general need to be drastically more **open**, **welcoming** and **supportive** towards everyone, especially underrepresented and marginalized groups
- **every single individual** can and will bring something important to the scientific table, **if they get the chance to**





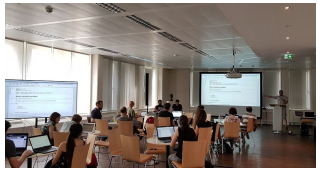

* photos from private collection, depicting various hackathons

FAIR data through community-driven development of standards and beyond

*adapted from Elizabeth DuPre

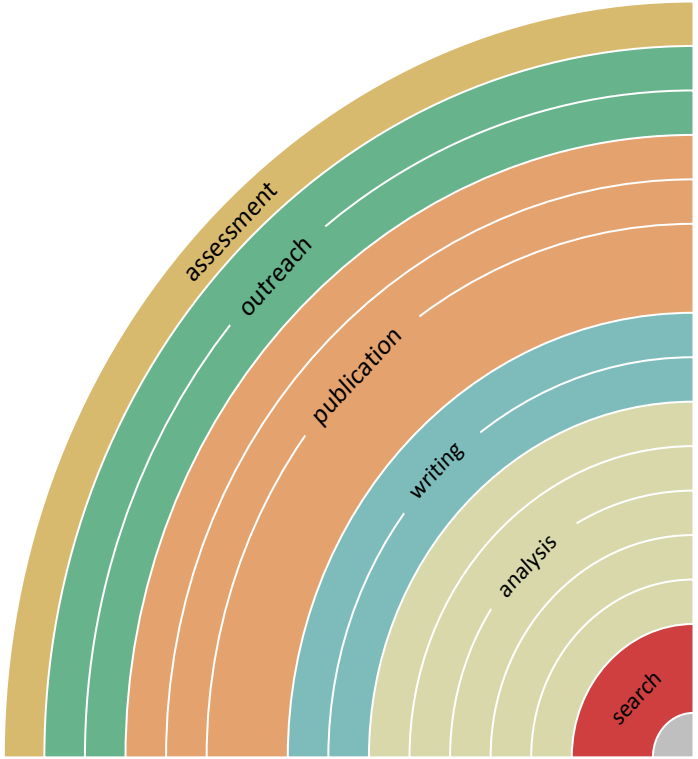
- **combine experience** across all levels (“big picture” folks and those that actually work with data)
- **community is required** as no one has all use cases
- **open discussion and development** necessary for acceptance and adaptation
- **community-driven standards** enable new kinds of science
- **everything needed for FAIR research outcomes exist** but individuals need to be empowered to use them and supported along the way






* photos from private collection, depicting various hackathons

FAIR data through community-driven development of standards and beyond

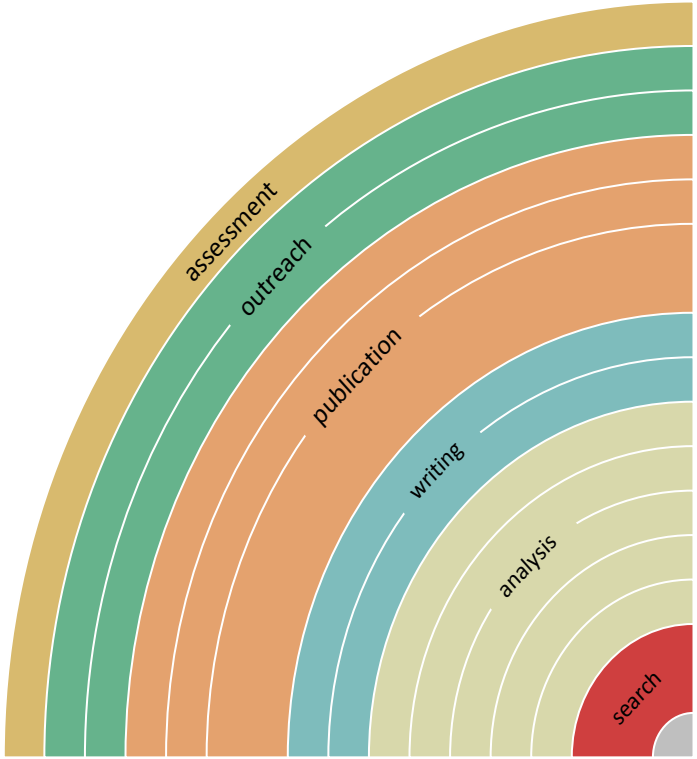


- adding alternative evaluation, e.g. with altmetrics
- communicating through social media, e.g. Twitter
- sharing posters & presentations, e.g. at FigShare
- using open licenses, e.g. CC0 or CC-BY
- publishing open access, ‘green’ or ‘gold’
- using open peer review, e.g. at journals or PubPeer
- sharing preprints, e.g. at OSF, arXiv or bioRxiv
- using actionable formats, e.g. with Jupyter or CoCalc
- open XML-drafting, e.g. at Overleaf or Authorea
- sharing protocols & workfl., e.g. at Protocols.io
- sharing notebooks, e.g. at OpenNotebookScience
- sharing code, e.g. at GitHub with GNU/MIT license
- sharing data, e.g. at Dryad, Zenodo or Dataverse
- pre-registering, e.g. at OSF or AsPredicted
- commenting openly, e.g. with Hypothes.is
- using shared reference libraries, e.g. with Zotero
- sharing (grant) proposals, e.g. at RIO

Bianca Kramer & Jeroen Bosman https://101innovations.wordpress.com

DOI: 10.5281/zenodo.1147025

FAIR data through community-driven development of standards and beyond



- necessities - support
- prerequisites - initiatives

- necessities - repositories (standardized)
- necessities - more than a PDF

- necessities - reporting (standardized)

- necessities - documentation
- necessities - virtualization (standardized)
- necessities - workflows (standardized)
- necessities - validation (standardized)
- necessities - standardization
- prerequisites - data sharing (standardized)
- prerequisites - preregistration (standardized)

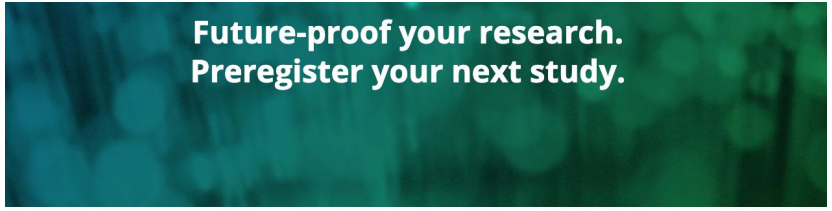
DOI: 10.5281/zenodo.1147025

FAIR data through community-driven development of standards and beyond

- necessities - standardization
- prerequisites - data sharing
- **prerequisites - preregistration**

DOI: 10.5281/zenodo.1147025

FAIR prerequisites - preregistration (standardized)



What is Preregistration?

When you preregister your research, you're simply specifying your research plan in advance of your study and submitting it to a registry.

Preregistration separates *hypothesis-generating* (exploratory) from *hypothesis-testing* (confirmatory) research. Both are important. But the same data cannot be used to generate *and* test a hypothesis, which can happen unintentionally and reduce the credibility of your results. Addressing this problem through planning improves the quality and transparency of your research. This helps you clearly report your study and helps others who may wish to build on it.

For additional insight and context, you can read The Preregistration Revolution. (preprint)

Get Started Now

Confirmatory Research

- Hypothesis testing
- Results are held to the highest standards
- Data-independent
- Minimizes false positives
- P-values retain diagnostic value
- Inferences may be drawn to wider population

Exploratory Research

- Hypothesis generating
- Results deserve to be replicated and confirmed
- Data-dependent
- Minimizes false negatives in order to find unexpected discoveries
- P-values lose diagnostic value
- Not useful for making inferences to any wider population

Create a new pre-registration CREATE

Just trying it out; make this pre-registration self-destroy in 24 hours.

See your pre-registrations (e.g., to share with reviewers or make public)

email address you have used in AsPredicted

Sign in

I cannot access my AsPredicted email account anymore

WHAT IS ASPREDICTED?

AsPredicted is a platform that makes it easy for researchers to pre-register their studies, and easy for others to read and evaluate those pre-registrations. To pre-register a study on AsPredicted, a researcher answers nine simple questions about their research design and analyses. The platform then generates a time-stamped, single page .pdf document that includes a unique URL for verification.

HOW DOES IT WORK?

- One author creates the pre-registration.
- Participating authors are emailed, requesting approval.
- If all approve, it is saved but remains private until an author makes it public; or remains private forever. (Why?)
- Authors may share an anonymous version of the pre-registration with reviewers.
- If made public, the final .pdf (sample) is automatically stored in the web-archive.

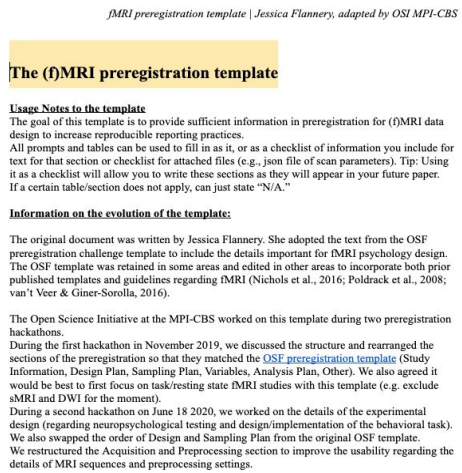
WHAT IF THINGS DON'T GO 'AS PREDICTED'?

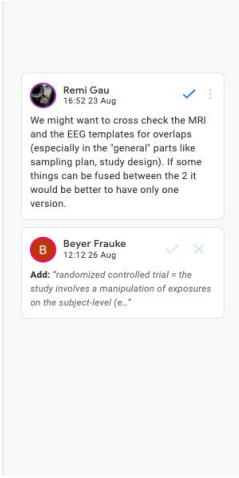
You can just say so in the paper:

- 'Contrary to expectations, we found that...'
- 'Unexpectedly, we also found that...'
- 'In addition to the analyses we pre-registered we also ran...'
- 'We encountered an unexpected situation, and followed our Standard Operating Procedure'(.pdf)

FAIR prerequisites - preregistration (standardized)

https://www.cbs.mpg.de/1543216/preregistration-hackathon





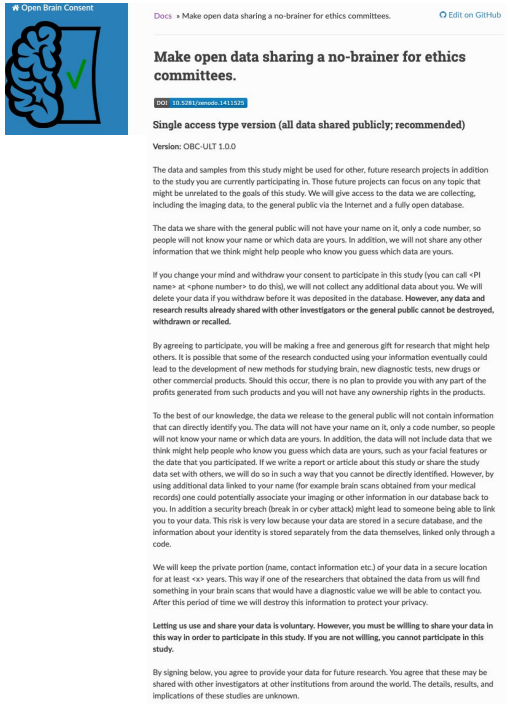
FAIR data through community-driven development of standards and beyond

- necessities - standardization
- **prerequisites - data sharing**
- prerequisites - preregistration

DOI: 10.5281/zenodo.1147025

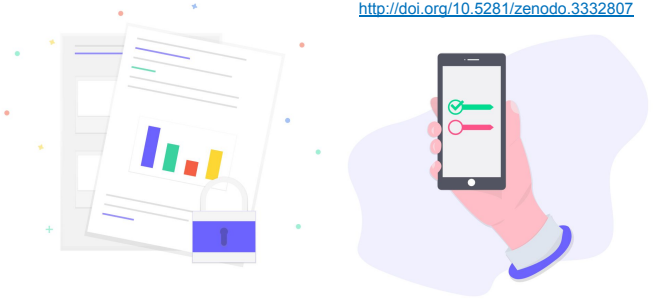
FAIR prerequisites - data sharing (standardized)

https://open-brain-consent.readthedocs.io/



- data privacy & rights > FAIR/open science
- data privacy & rights need to be protected
- identification attacks & data breaches
- complicated legal issues and varying interpretations of laws
- researchers not trained

http://doi.org/10.5281/zenodo.3332807



FAIR prerequisites - data sharing (standardized)

https://open-brain-consent.readthedocs.io/

Docs » Make open data sharing a no-brainer for ethics committees. Edit on GitHub

Make open data sharing a no-brainer for ethics committees.

DOI 10.5281/zenodo.14115175

Single access type version (all data shared publicly; recommended)

Version: OBC-ULT 1.0.0

The data and samples from this study might be used for other, future research projects in addition to the study you are currently participating in. Those future projects can focus on any topic that might be unrelated to the goals of this study. We will give access to the data we are collecting, including the imaging data, to the general public via the Internet and a fully open database.

The data we share with the general public will not have your name on it, only a code number, so people will not know your name or which data are yours. In addition, we will not share any other information that we think might help people who know you guess which data are yours.

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.**

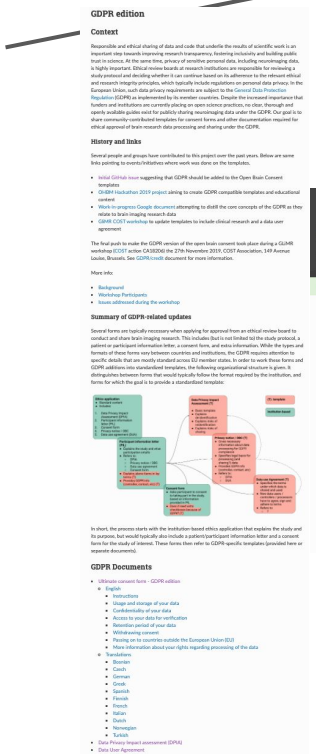
By agreeing to participate, you will be making a free and generous gift for research that might help others. It is possible that some of the research conducted using your information eventually could lead to the development of new methods for studying brain, new diagnostic tests, new drugs or other commercial products. Should this occur, there is no plan to provide you with any part of the profits generated from such products and you will not have any ownership rights in the products.

To the best of our knowledge, the data we release to the general public will not contain information that can directly identify you. The data will not have your name on it, only a code number, so people will not know your name or which data are yours. In addition, the data will not include data that we think might help people who know you guess which data are yours, such as your facial features or the date that you participated. If we write a report or article about this study or share the study data set with others, we will do so in such a way that you cannot be directly identified. However, by using additional data linked to your name (for example brain scans obtained from your medical records) one could potentially associate your imaging or other information in our database back to you. In addition a security breach (break in or cyber attack) might lead to someone being able to link you to your data. This risk is very low because your data are stored in a secure database, and the information about your identity is stored separately from the data themselves, linked only through a code.

We will keep the private portion (name, contact information etc.) of your data in a secure location for at least <x> years. This way if one of the researchers that obtained the data from us will find something in your brain scans that would have a diagnostic value we will be able to contact you. After this period of time we will destroy this information to protect your privacy.

Letting us use and share your data is voluntary. However, you must be willing to share your data in this way in order to participate in this study. If you are not willing, you cannot participate in this study.

By signing below, you agree to provide your data for future research. You agree that these may be shared with other investigators at other institutions from around the world. The details, results, and implications of these studies are unknown.



GDPR edition

Context

Responsible and ethical sharing of data and code that underlie the results of scientific work is an important step towards improving research transparency, fostering inclusivity and building public trust in science. At the same time, privacy of sensitive personal data, including neuroimaging data, is highly important. Ethical review boards at research institutions are responsible for reviewing a study protocol and deciding whether it can continue based on its adherence to the relevant ethical and research integrity principles, which typically include regulations on personal data privacy. In the European Union, such data privacy requirements are subject to the General Data Protection Regulation (GDPR) as implemented by its member countries. Despite the increased importance that funders and institutions are currently placing on open science practices, no clear, thorough and openly available guides exist for publicly-sharing neuroimaging data under the GDPR. Our goal is to share community-contributed templates for consent forms and other documentation required for ethical approval of brain research data processing and sharing under the GDPR.

History and links

Several people and groups have contributed to this project over the past years. Below are some links pointing to events/initiatives where work was done on the templates.

- Initial GitHub issue suggesting that GDPR should be added to the Open Brain Consent templates
- OHBM Hackathon 2019 project aiming to create GDPR compatible templates and educational content
- Work-in-progress Google document attempting to distill the core concepts of the GDPR as they relate to brain imaging research data
- GDPR COST workshop to update templates to include clinical research and a data user agreement

The final push to make the GDPR version of the open brain consent took place during a GLIMR workshop (COST action CA18236) the 27th November 2019. COST Association, 149 Avenue Louise, Brussels. See GDPR.tex46 document for more information.

More info:

- Background
- Workshop Participants
- Issues addressed during the workshop

Summary of GDPR-related updates

Several forms are typically necessary when applying for approval from an ethical review board to conduct and share brain imaging research. This includes (but is not limited to) the study protocol, a patient or participant information letter, a consent form, and extra information. While the types and formats of these forms vary between countries and institutions, the GDPR requires attention to specific details that are mostly standard across EU member states. In order to work these forms and GDPR additions into standardized templates, the following organizational structure is given. It distinguishes between forms that would typically follow the format required by the institution, and forms for which the goal is to provide a standardized template:

In short, the process starts with the institution-based ethics application that explains the study and its purpose, but would typically also include a patient/participant information letter and a consent form for the study of interest. These forms then refer to GDPR-specific templates (provided here or separate documents).

GDPR Documents

- Ultimate consent form - GDPR edition
  - English
    - Instructions
    - Usage and storage of your data
    - Confidentiality of your data
    - Access to your data for verification
    - Retention period of your data
    - Withdrawing consent
    - Passing on to countries outside the European Union (EU)
    - More information about your rights regarding processing of the data
  - Translations
    - Bosnian
    - Czech
    - German
    - Greek
    - Spanish
    - Finnish
    - French
    - Italian
    - Dutch
    - Norwegian
    - Turkish
- Data Privacy Impact assessment (DPIA)
- Data User Agreement

2018

2018

update consent form new EU GDPR data law #24

Open PeerHerholz opened this issue on Jun 1, 2018 · 13 comments

PeerHerholz commented on Jun 1, 2018 Member

Hey everyone,

I guess most of have heard that the EU made some changes regarding data protection and you got one or the other "we've updated our privacy policy" mails.
It's called the General Data Protection Regulation (GDPR) and you can find information about it here and here.

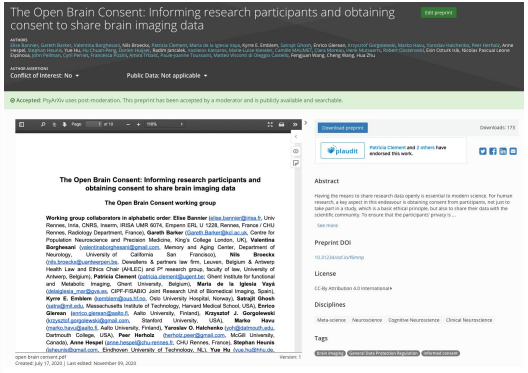
As I recently put some work into making the open brain consent the standard data sharing form here I had "some" discussions, also regarding the GDPR. Based on that, I wanted to ask if you think the form should be updated including add ons related to the GDPR? Maybe an GDPR version for european folks/studies?
After going through the legal text and in accordance with our local ethics committee, adding the contact information of the state's data protection officer and the person that's responsible for data analysis is sufficient so far (our updated version was already accepted here).

Would be great to get some ideas and start discussions.

Best, Peer

2019

2020



The Open Brain Consent: Informing research participants and obtaining consent to share brain imaging data

Edit preprint

Author Additions

Conflict of Interest: No

Public Data: Not applicable

Accepted: PsyArXiv uses post-moderation. This preprint has been accepted by a moderator and is publicly available and searchable.

The Open Brain Consent: Informing research participants and obtaining consent to share brain imaging data

The Open Brain Consent working group

Working group collaborators in alphabetic order: Elise Bannier (elise.bannier@inria.fr; Univ Rennes, Inria, CNRS, Inserm, IRISA UMR 6074, Empenn ERL U 1228, Rennes, France / CHU Rennes, Radiology Department, France), Gareth Barker (Gareth.Barker@kcl.ac.uk; Centre for Population Neuroscience and Precision Medicine, King's College London, UK), Valentina Borghesani (valentinaborghesani@gmail.com; Memory and Aging Center, Department of Neurology, University of California San Francisco), Nils Broeckx (nils.broeckx@cannesgroup.be; Dewallens & partners law firm, Leuven, Belgium & Antwerp Health Law and Ethics Chair (AHLEC) and PI research group, faculty of law, University of Antwerp, Belgium), Patricia Clement (patricia.clement@ugent.be; Ghent Institute for Functional and Metabolic Imaging, Ghent University, Belgium), Maria de la Iglesia Vaya (delaiglesia_mar@gva.es; CIPF-FISABIO Joint Research Unit of Biomedical Imaging, Spain), Kyrre E. Emblem (kemblem@mgh.harvard.edu; Oslo University Hospital, Norway), Satrajit Ghosh (satra@mit.edu; Massachusetts Institute of Technology, Harvard Medical School, USA), Enrico Glerean (enrico.glerean@aalto.fi; Aalto University, Finland), Krzysztof J. Gorgolewski (krzysztof.gorgolewski@gmail.com; Stanford University, USA), Marko Havu (marko.havu@aalto.fi; Aalto University, Finland), Yaroslav O. Halchenko (yoh@dartmouth.edu; Dartmouth College, USA), Peer Herholz (herholz.peer@gmail.com; McGill University, Canada), Anne Hospel (anne.hospel@univ-rennes1; CHU Rennes, France), Stephan Heunis (stheunis@gmail.com; Eindhoven University of Technology, NL), Yue Hu (yue.hu@hhu.de;

Preprint DOI

10.31234/osf.io/fkmyg

License

CC-By Attribution 4.0 International

Disciplines

Meta-science · Neuroscience · Cognitive Neuroscience · Clinical Neuroscience

Tags

Brain Imaging · General Data Protection Regulation · Informed consent

open brain consent.pdf

Created: July 17, 2020 | Last edited: November 09, 2020

Version: 1

- GDPR edition, accepted example submissions, DUAs (in over 10 languages)
- preprint with 35 coauthors from all over the world

2019

McGill UNIVERSITY · neuro · Brams International Laboratory for Brain, Music, and Sound Research

FAIR data through community-driven development of standards and beyond

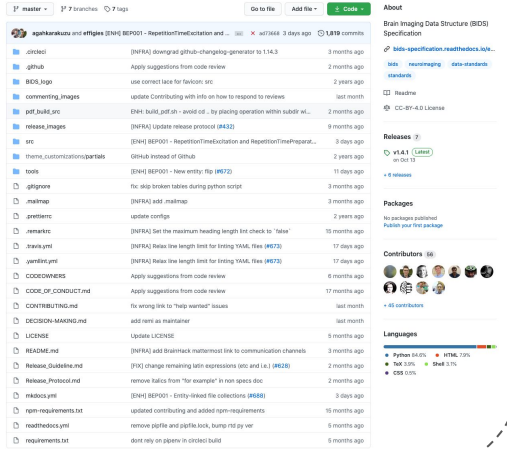
- **necessities - standardization**
- prerequisites - data sharing
- prerequisites - preregistration

DOI: 10.5281/zenodo.1147025

FAIR necessities - standardization - fantastic standards in the wild

BIDS
BRAIN IMAGING DATA STRUCTURE

120+ contributors*



A specification for psychological datasets

PSYCH-DS

20+ contributors*

Psych-DS

A technical specification for psychological datasets

Version 0.3.0 ("POST SIPS2020" DRAFT) - IN PROGRESS

Available under the CC-BY 4.0 International license
Based on the BIDS specification for fMRI data, with inspiration from the BIDS Eyetracking extension draft.

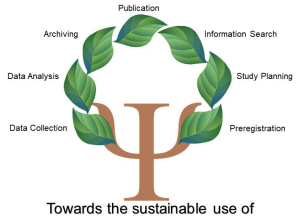
Welcome! If this is your first time checking out this project (and especially if this is your first time working with a technical specification document), please also check out the project landing page for an explanation of the goals of this project and notes on how you can contribute.

If you are working on creating a Psych-DS compliant dataset, please check out the How-To page!

The technical specification document should focus closely on the definition of Psych-DS. Background information on how or why to use Psych-DS (or comments for tool implementors) should be factored either into the How-To page or into the position paper draft.

This document contains a draft of a data structure standard for psychological datasets. It is a community effort to define standards in data / metadata. This is a working document in draft stage and ***comments are welcome***.

This specification is an extension of BIDS, and general principles are shared. The Psych-DS specification aims to work for many different settings and facilitate integration with other types of data.



Towards the sustainable use of psychological research data

The Project PsyCuraDat - Development of user-oriented curation criteria for psychological research data

Project description

In the context of the omnipresent Open Science Movement, all stakeholders in psychological research (i.e. researchers, journals, and funders) increasingly strive for more transparency in order to foster scientific exchange and the quality of scientific work. However, to reach this overarching goal of the Open Science Movement, there is a need for discipline-specific documentation standards not only for research articles, but also for the data representing the basis of those articles. Against this background, the project „PsyCuraDat: Development of user-oriented curation criteria for psychological research data", which is funded by the Federal Ministry of Education and Research (BMBF), pursues two strategic objectives:

1. To meet researchers' needs in their role as data users, the project aims at developing a documentation standard for psychological research data considering the discipline specific methods relevant for reusing these data.
2. To meet researchers' needs in their role as contributors of research data, the project aims at developing a documentation standard considering determinants defined by the research process as well as by funders and journals.

Working towards these two strategic goals, PsyCuraDat generally aims at promoting research economy and research integrity in Psychology by enabling a more effective and efficient documentation and reuse of psychological research data.

McGill UNIVERSITY neuro Brams International Laboratory for Brain, Music, and Sound Research

*only some are shown

FAIR necessities - standardization - the idea of BIDS

How do you manage your data?

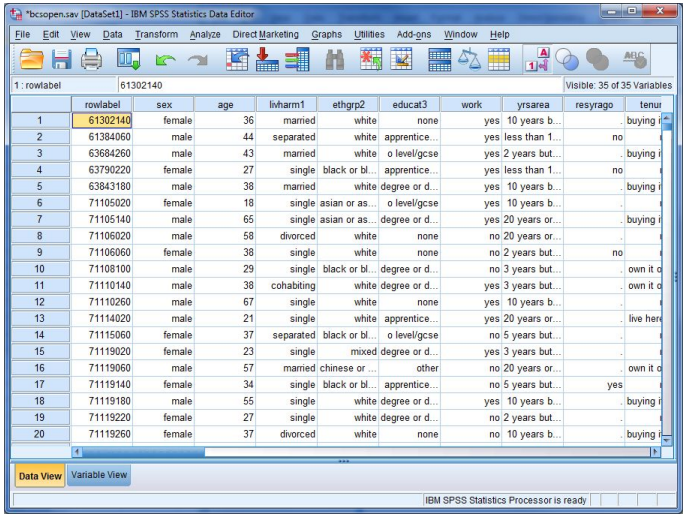
- storage, structure, metadata, version control?

How do you share your data?

- colleagues, students, other researchers?

The Problem with heterogeneity in data management is it is hard for others (and you) to understand your data and keep track of changes

- unnecessary metadata input
- codes / scripts have to be adapted
- huge effort to automate workflows and no way to automatically validate data sets
- sharing data becomes a hassle



https://bizlib247.files.wordpress.com/2014/04/data-editor-labels.png

FAIR necessities - standardization - the idea of BIDS

* content liberally borrowed from
  Guiomar Niso's [BIDS ecosystem presentation]

FAIR necessities - standardization - the idea of BIDS … *with Marie Kondo*

- dicomdir/
  - 1208200617178_22/
    - 1208200617178_22_8973.dcm
    - 1208200617178_22_8943.dcm
    - 1208200617178_22_2973.dcm
    - 1208200617178_22_8923.dcm
    - 1208200617178_22_4473.dcm
    - 1208200617178_22_8783.dcm
    - 1208200617178_22_7328.dcm
    - 1208200617178_22_9264.dcm
    - 1208200617178_22_9967.dcm
    - 1208200617178_22_3894.dcm
    - 1208200617178_22_3899.dcm
  - 1208200617178_23/
  - 1208200617178_24/
  - 1208200617178_25/

BIDS
BRAIN IMAGING DATA STRUCTURE

- my_dataset/
  - participants.tsv
  - sub-01/
    - anat/
      - sub-01_T1w.nii.gz
    - func/
      - sub-01_task-rest_bold.nii.gz
      - sub-01_task-rest_bold.json
    - dwi/
      - sub-01_dwi.nii.gz
      - sub-01_dwi.json
      - sub-01_dwi.bval
      - sub-01_dwi.bvec
  - sub-02/
  - sub-03/
  - sub-04/

FAIR necessities - standardization - the idea of BIDS

* content liberally borrowed from
Chris Markiewicz's BIDS ecosystem presentation

https://bids-specification.readthedocs.io/en/stable/
https://bids.neuroimaging.io/

- **BIDS** is a **directory structure** and **naming convention**, based on common practices
- builds on **existing open standards** (`NIfTI`, `JSON`, `TSV`)
- intended for **human** and **machine legibility**
- **metadata** for all files and across different levels
- **maintained** and **extended** by a **clearly structured and open community**


```
sub-control01/
    anat/
        sub-control01_T1w.nii.gz
        sub-control01_T1w.json
        sub-control01_T2w.nii.gz
        sub-control01_T2w.json
    func/
        sub-control01_task-nback_bold.nii.gz
        sub-control01_task-nback_bold.json
        sub-control01_task-nback_events.tsv
        sub-control01_task-nback_physio.tsv.gz
        sub-control01_task-nback_physio.json
        sub-control01_task-nback_sbref.nii.gz
    dwi/
        sub-control01_dwi.nii.gz
        sub-control01_dwi.bval
        sub-control01_dwi.bvec
    fmap/
        sub-control01_phasediff.nii.gz
        sub-control01_phasediff.json
        sub-control01_magnitude1.nii.gz
        sub-control01_scans.tsv
code/
    deface.py
derivatives/
README
participants.tsv
dataset_description.json
CHANGES
```

```
{
  "test": {
    "LongName": "Education level",
    "Description": "Education level, self-rated by participant",
    "Levels": {
      "1": "Finished primary school",
      "2": "Finished secondary school",
      "3": "Student at university",
      "4": "Has degree from university"
    }
  },
  "bmi": {
    "LongName": "Body mass index",
    "Units": "kg/m^2",
    "TermURL": "http://purl.bioontology.org/ontology/SNOMEDCT/60621009"
  }
}
```

FAIR necessities - standardization - the idea of BIDS

* content liberally borrowed from
Guiomar Niso's BIDS ecosystem presentation

https://bids-specification.readthedocs.io/en/stable/
https://bids.neuroimaging.io/

dataset_description.json


```
{
"Name": "The mother of all experiments",
"BIDSVersion": "1.4.0",
"DatasetType": "raw",
"License": "CC0",
"Authors": [ "Paul Broca", "Carl Wernicke" ],
"Acknowledgements": "Special thanks to KB for help in formatting
this dataset in BIDS",
"HowToAcknowledge": "Please cite this paper: https://
www.ncbi.nlm.nih.gov/pubmed/0928",
"Funding": [ "National Institute of Neuroscience Grant F37823MFH1" ],
"EthicsApprovals": ["Human Research Protections Office (Protocol
AR0928" ],
"ReferencesAndLinks": [ "https://www.ncbi.nlm.nih.gov/pubmed/
0928", "Alzheimer et al (2015). Nature, 21. doi.org/0928" ],
"DatasetDOI": "10.0.2.3/dfjj.10"
}
```


BIDS
BRAIN IMAGING DATA STRUCTURE

example1/

- CHANGES
- README.tsv
- **dataset_description.json**
- participants.json
- participants.tsv
- **sub-001**
  - ses-01
  - ses-02
  - sub-001_scans.tsv
- **sub-002**
- **sub-003**

FAIR necessities - standardization - the idea of BIDS

* content liberally borrowed from
Guiomar Niso's BIDS ecosystem presentation

https://bids-specification.readthedocs.io/en/stable/
https://bids.neuroimaging.io/

participants.tsv

participant_id	age	sex
sub-001	34	M
sub-002	12	F
sub-003	33	F

participants.json

```
"age": {
    "Description": "age of the participant",
    "Units": "years"  },
"sex": {
    "Description": "sex as reported by the participant",
    "Levels": {
        "M": "male",
        "F": "female"  }
```

BIDS
BRAIN IMAGING DATA STRUCTURE

example1/

```
├── CHANGES
├── README.tsv
├── dataset_description.json
├── participants.json
├── participants.tsv
└── sub-001
    ├── ses-01
    ├── ses-02
    └── sub-001_scans.tsv
├── sub-002
└── sub-003
```

FAIR necessities - standardization - the idea of BIDS

* content liberally borrowed from
Guiomar Niso's BIDS ecosystem presentation

https://bids-specification.readthedocs.io/en/stable/
https://bids.neuroimaging.io/

sub-001_ses-02_task-rest_eeg.json


```
{
"InstitutionName": "Montreal Neurological Institute",
"InstitutionAddress": "3801 University, Montreal QC",
"Manufacturer": "Brain Products",
"ManufacturersModelName": "BrainAmp DC",
"DeviceSerialNumber": "11035",
"PowerLineFrequency": 60,
"SamplingFrequency": 2400,
"EEGPlacementScheme":"10 percent system",
"EEGReference": "left mastoid",
"SoftwareFilters": {"Anti-aliasing filter": {"half-amplitude cutoff (Hz)": 500, "Roll-off": "6dB/Octave"}},
"CapManufacturer": "EasyCap",
"CapManufacturersModelName": "M1-ext",
"EOGChannelCount": 2,
"ECGChannelCount": 1,
"EMGChannelCount": 0,
"RecordingDuration": 600,"RecordingType": "continuous",
"EpochLength": 0,
...
```


BIDS
BRAIN IMAGING DATA STRUCTURE

```
sub-001
├── ses-02
│   ├── meg
│   │   ├── sub-001_ses-02_coordsystem.json
│   │   ├── sub-001_ses-02_headshape.pos
│   │   ├── sub-001_ses-02_task-rest_run-01_channels.tsv
│   │   ├── sub-001_ses-02_task-rest_run-01_meg.ds
│   │   └── sub-001_ses-02_task-rest_run-01_meg.json
├── ses-03
│   ├── eeg
│   │   ├── sub-001_ses-03_task-rest_run-01_eeg.edf
│   │   └── sub-001_ses-03_task-rest_run-01_eeg.json
│   ├── beh
│   │   └── sub-001_ses-03_task-oddball_run-01_events.tsv
└── sub-001_scans.tsv
```

FAIR necessities - standardization - the idea of BIDS

* content liberally borrowed from
Guiomar Niso's BIDS ecosystem presentation

https://bids-specification.readthedocs.io/en/stable/
https://bids.neuroimaging.io/

- **adoption is crucial**
  - minimize complexity and facilitate adoption
  - don't reinvent the wheel
  - reuse existing methods and technologies

- **80/20** rule
  - tackle 80% of most commonly used cases

- **engagement of a global community**
  - everyone can provide input and work on maintenance and extension

BIDS
BRAIN IMAGING DATA STRUCTURE

```
sub-control01/
    anat/
        sub-control01_T1w.nii.gz
        sub-control01_T1w.json
        sub-control01_T2w.nii.gz
        sub-control01_T2w.json
    func/
        sub-control01_task-nback_bold.nii.gz
        sub-control01_task-nback_bold.json
        sub-control01_task-nback_events.tsv
        sub-control01_task-nback_physio.tsv.gz
        sub-control01_task-nback_physio.json
        sub-control01_task-nback_sbref.nii.gz
    dwi/
        sub-control01_dwi.nii.gz
        sub-control01_dwi.bval
        sub-control01_dwi.bvec
    fmap/
        sub-control01_phasediff.nii.gz
        sub-control01_phasediff.json
        sub-control01_magnitude1.nii.gz
        sub-control01_scans.tsv
code/
    deface.py
derivatives/
README
participants.tsv
dataset_description.json
CHANGES
```

```
{
  "test": {
    "LongName": "Education level",
    "Description": "Education level, self-rated by participant",
    "Levels": {
      "1": "Finished primary school",
      "2": "Finished secondary school",
      "3": "Student at university",
      "4": "Has degree from university"
    }
  },
  "bmi": {
    "LongName": "Body mass index",
    "Units": "kg/m^2",
    "TermURL": "http://purl.bioontology.org/ontology/SNOMEDCT/60621009"
  }
}
```

FAIR necessities - standardization - the idea of BIDS

https://bids-specification.readthedocs.io/en/stable/
https://bids.neuroimaging.io/

* content liberally borrowed from
Guiomar Niso's BIDS ecosystem presentation

- BIDS is constantly expanded & updated through community driven development
    - Positron Emission Tomography (PET)
    - Common Derivatives
    - Models Specification
    - Magnetoencephalography (MEG)
    - Electroencephalography (EEG)
    - intracranial Electroencephalography (iEEG)
    - Eye Tracking including Gaze Position and Pupil Size
    - Susceptibility Weighted Imaging (SWI)
    - Genetic information
    - Microelectrode Recordings (MER)
    - ….

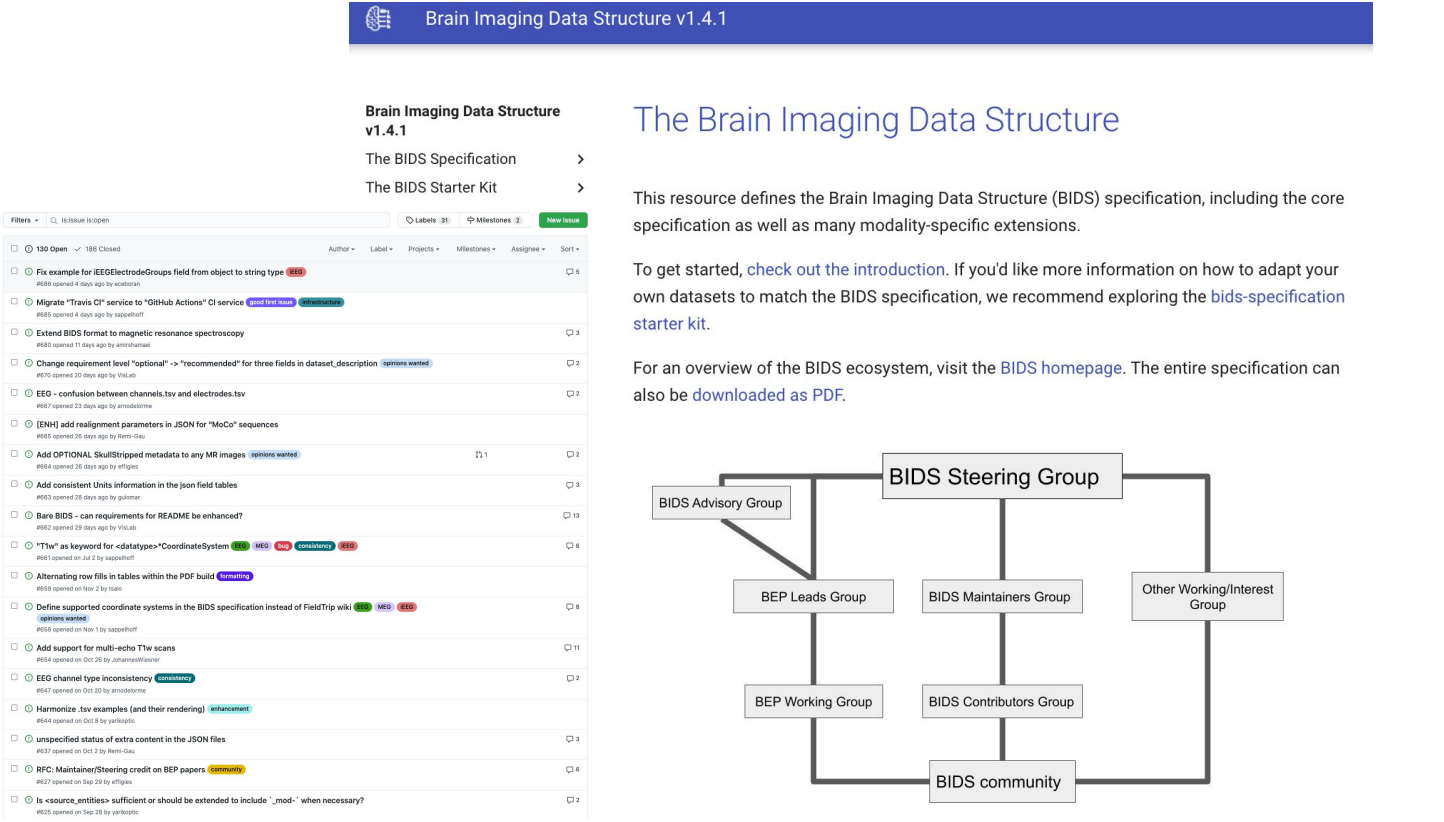
BEP	Title	Moderators/Leads
BEP001	Structural acquisitions with multiple contrasts (multi echo, flip angle, inversion time) seqs	G. de Hollander, K. Whitaker
BEP002	BIDS Models	T. Yarkoni
BEP004	Susceptibility Weighted Imaging (SWI)	Available
BEP005	**Arterial Spin Labeling (ASL)**	**Mutsaerts, Clement, Petr, Castellaro**
BEP009	**Positron Emission Tomography (PET)**	**M. Ganz**
BEP011	Structural preprocessing derivatives	V. Siless
BEP012	Functional preprocessing derivatives	C.Maumet, Markiewicz
BEP013	Resting state fMRI derivatives	S. Giavasis
BEP014	Affine transformations and nonlinear field warps	O. Esteban
BEP016	Diffusion weighted imaging derivatives	F. Pestilli, O. Esteban
BEP017	Generic BIDS connectivity data schema	E. Duff, P. McCarthy
BEP019	DICOM Metadata	S. Ghosh

BEP	Title	Moderators/Leads
BEP020	Eye Tracking including Gaze Position and Pupil Size (ET)	B. Gagl, D. Draschkow
BEP021	Common Electrophysiological Derivatives	M.Jas, Appelhoff, Pernet, G.Niso, R.Oostenveld
BEP022	Magnetic Resonance Spectroscopy (MRS)	M. Mikkelsen, W. Clarke
BEP023	PET Preprocessing derivatives	M.Noergaard, Searle, Ganz
BEP024	Computed Tomography scan (CT)	H. Boniface
BEP025	Medical Imaging Data structure (MIDS)	J.M.Saborit, M.de la Iglesia
BEP026	Microelectrode Recordings (MER)	G. Gilmore
BEP027	BIDS Execution	C. Markiewicz, G. Kiar
BEP028	Provenance	S. Ghosh, C. Maumet
BEP029	Virtual and physical motion data	S. Jeung, J. Welzel
BEP030	Near Infrared Spectroscopy (NIRS)	R. Luke, L. Pollonini
BEP031	Microscopy	M.Bourget, J.Cohen-Adad

BEP	Title	Moderators/Leads
BEP003	Common Derivatives	C. Markiewicz
BEP006	Electroencephalography (EEG)	C. Pernet, S. Appelhoff, R. Oostenveld
BEP007	Hierarchical Event Descriptor (HED) Tags	C. Gorgolewski
BEP008	Magnetoencephalography (MEG)	G. Niso
BEP010	intracranial Electroencephalography (iEEG)	C. Holdgraf, D. Hermes
BEP018	Genetic information	C. Pernet, C. Moreau, T. Nichols

FAIR necessities - standardization - the idea of BIDS

https://bids-specification.readthedocs.io/en/stable/
https://bids.neuroimaging.io/



FAIR necessities - standardization - the idea of BIDS

https://bids-specification.readthedocs.io/en/stable/
https://bids.neuroimaging.io/

* content liberally borrowed from
  Guiomar Niso's BIDS ecosystem presentation

BIDS Steering Group 2019-2020

Guiomar Niso
Spain
MEG

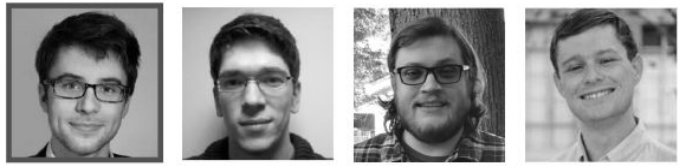
Melanie Ganz
Denmark
PET

Robert Oostenveld
The Netherlands
M/i/EEG

Russ Poldrack
USA
fMRI

Kirstie Whitaker
UK
MRI

BIDS Maintainers

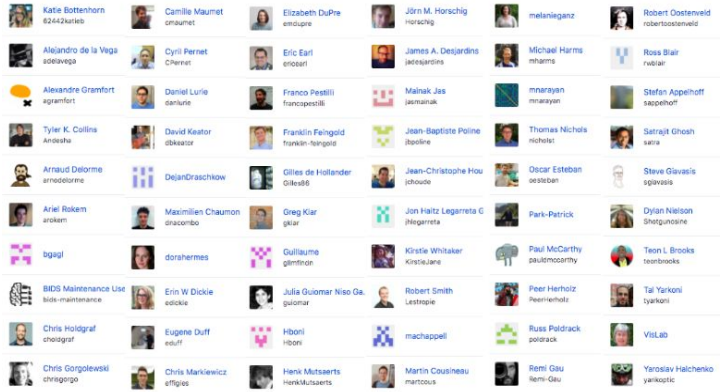


Stefan Appelhoff
Germany

Chris Markiewicz
USA

Taylor Salo
USA

Franklin Feingold
USA

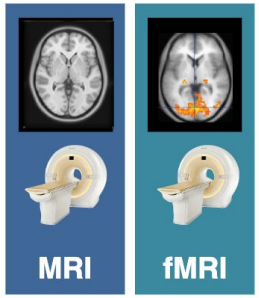


*only some contributors are shown

FAIR necessities - standardization - the idea of BIDS

https://bids-specification.readthedocs.io/en/stable/
https://bids.neuroimaging.io/

* content liberally borrowed from
  Guiomar Niso's BIDS ecosystem presentation



SCIENTIFIC DATA

OPEN

SUBJECT CATEGORIES
» Data publication and archiving
» Research data

The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments

Received: 18 December 2015
Accepted: 19 May 2016
Published: 21 June 2016

Krzysztof J. Gorgolewski[1], Tibor Auer[2], Vince D. Calhoun[3,4], R. Cameron Craddock[5,6], Samir Das[7], Eugene P. Duff[8], Guillaume Flandin[9], Satrajit S. Ghosh[10,11], Tristan Glatard[7,12], Yaroslav O. Halchenko[13], Daniel A. Handwerker[14], Michael Hanke[15,16], David Keator[17], Xiangrui Li[18], Zachary Michael[19], Camille Maumet[20], B. Nolan Nichols[21,22], Thomas E. Nichols[20,23], John Pellman[9], Jean-Baptiste Poline[24], Ariel Rokem[25], Gunnar Schaefer[1,26], Vanessa Sochat[27], William Triplett[1], Jessica A. Turner[1,28], Gaël Varoquaux[29] & Russell A. Poldrack[1]

(Gorgolewski et al.2016)

SCIENTIFIC DATA

OPEN

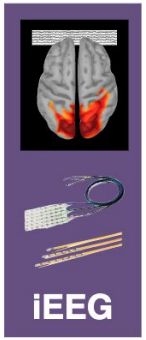
COMMENT

EEG-BIDS, an extension to the brain imaging data structure for electroencephalography

Received: 16 January 2019
Accepted: 7 May 2019
Published online: 25 June 2019

Cyril R. Pernet[1], Stefan Appelhoff[2], Krzysztof J. Gorgolewski[3], Guillaume Flandin[4], Christophe Phillips[5], Arnaud Delorme[6,7] & Robert Oostenveld[8,9]

(Pernet et al., 2019)



SCIENTIFIC DATA

OPEN

COMMENT

iEEG-BIDS, extending the Brain Imaging Data Structure specification to human intracranial electrophysiology

Received: 29 January 2019
Accepted: 24 May 2019
Published online: 25 June 2019

Christopher Holdgraf[1-16], Stefan Appelhoff[2], Stephan Bickel[3], Kristofer Bouchard[4], Sasha D'Ambrosio[5], Olivier David[6], Orrin Devinsky[7], Benjamin Dichter[8], Adeen Flinker[7], Brett L. Foster[9], Krzysztof J. Gorgolewski[8], Iris Groen[10], David Groppe[11], Aysegul Gunduz[12], Liberty Hamilton[13], Christopher J. Honey[14], Mainak Jas[15], Robert Knight[16], Jean-Philippe Lachaux[17], Jonathan C. Lau[18], Christopher Lee-Messer[8], Brian N. Lundstrom[19], Kai J. Miller[20], Jeffrey G. Ojemann[21], Robert Oostenveld[22], Natalia Petridou[23], Gio Piantoni[24], Andrea Pigorini[5], Nader Pouratian[25], Nick F. Ramsey[26], Arjen Stolk[16], Nicole C. Swann[28], François Tadel[6,27], Bradley Voytek[28], Brian A. Wandell[8], Jonathan Winawer[10], Kirstie Whitaker[29,32], Lyuba Zehl[30] & Dora Hermes[8,24,31]

(Holdgraf et al., 2019)

SCIENTIFIC DATA

OPEN

Comment: MEG-BIDS, the brain imaging data structure extended to magnetoencephalography

Received: 14 November 2017
Accepted: 3 May 2018
Published: 19 June 2018

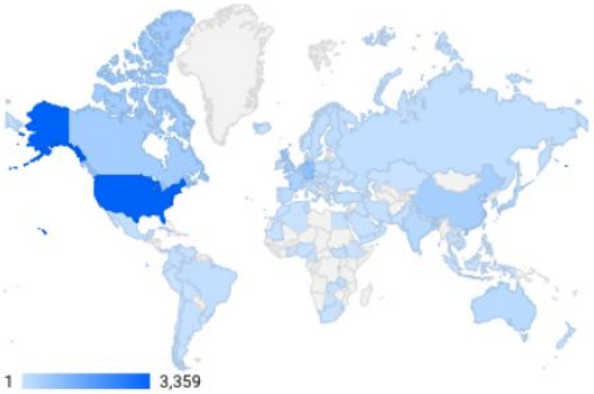
Guiomar Niso[1,2], Krzysztof J. Gorgolewski[3], Elizabeth Bock[1], Teon L. Brooks[3], Guillaume Flandin[3], Alexandre Gramfort[5,6], Richard N. Henson[7], Mainak Jas[5], Vladimir Litvak[4], Jeremy T. Moreau[1], Robert Oostenveld[8,9], Jan-Mathijs Schoffelen[8], Francois Tadel[1,10,11], Joseph Wexler[1] & Sylvain Baillet[1]

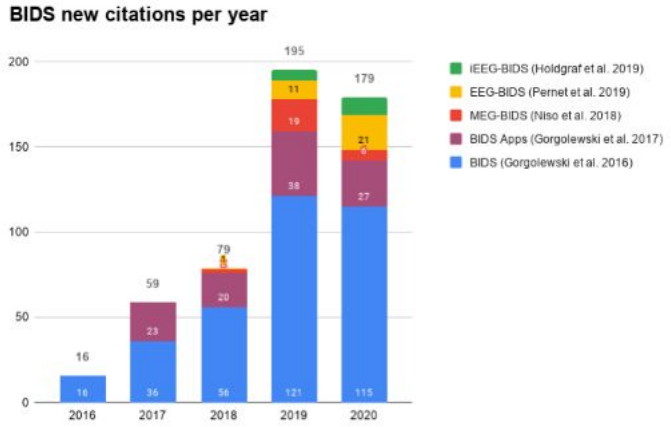
(Niso et al., 2018)

FAIR necessities - standardization - the idea of BIDS

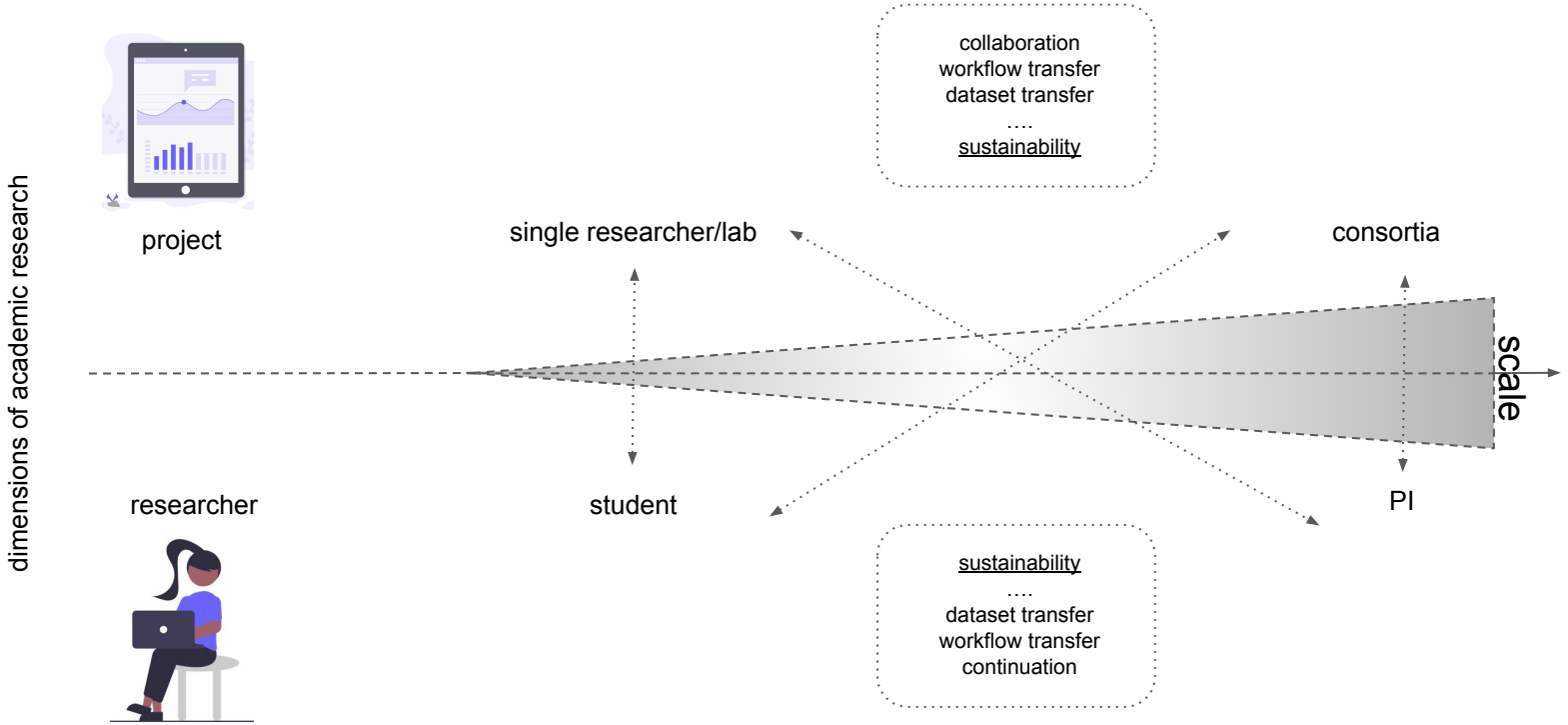
https://bids-specification.readthedocs.io/en/stable/
https://bids.neuroimaging.io/

* content liberally borrowed from
  Guiomar Niso's BIDS ecosystem presentation

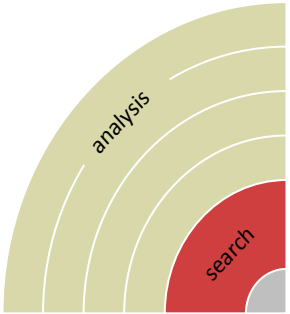




FAIR necessities - standardization - the benefits of BIDS



FAIR data through community-driven development of standards and beyond



- necessities - documentation
- **necessities - virtualization (standardized)**
- **necessities - workflows (standardized)**
- **necessities - validation (standardized)**
- necessities - standardization
- prerequisites - data sharing (standardized)
- prerequisites - preregistration (standardized)

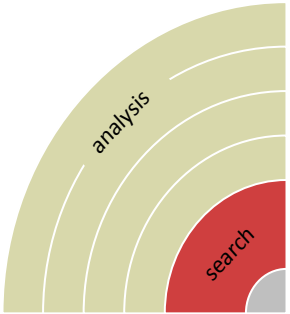
BIDS Apps

jupyter Neuroscout

BIDS BRAIN IMAGING DATA STRUCTURE

DOI: 10.5281/zenodo.1147025

FAIR data through community-driven development of standards and beyond



- necessities - documentation
- necessities - virtualization (standardized)
- necessities - workflows (standardized)
- **necessities - validation (standardized)**
- necessities - standardization
- prerequisites - data sharing (standardized)
- prerequisites - preregistration (standardized)

DOI: 10.5281/zenodo.1147025

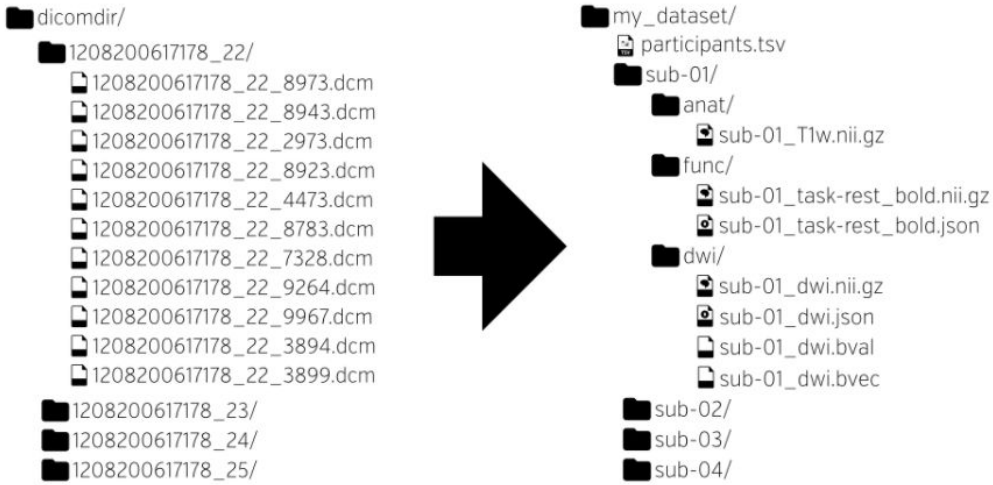
FAIR necessities - standardization - the benefits of BIDS - data conversion

* content liberally borrowed from
  Chris Markiewicz's BIDS ecosystem presentation

https://bids-specification.readthedocs.io/en/stable/
https://bids.neuroimaging.io/

BIDS
BRAIN IMAGING DATA STRUCTURE

- vast amount of converters to convert raw data to BIDS in a (semi) automated and reproducible manner
- support for various modalities and input data types
- AFNI BIDS-tools, BIDS2ISATab, BIDSto3col, BIDS2NDA, bidsfy, bidskit, dac2bids, Data2Bids, Dcm2bids, heudiconv, DCM2NIIx, ReproIn, XNAT2BIDS, bids2xar, BIDS2NIDM, BIDScoin, MNE-BIDS...

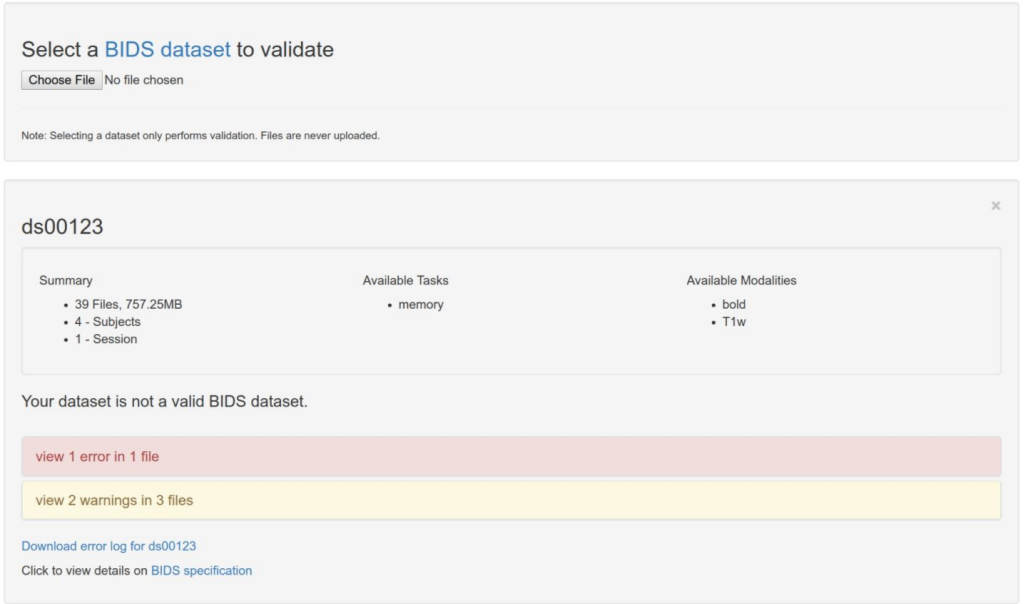


FAIR necessities - standardization - the benefits of BIDS - data validation

* content liberally borrowed from
Chris Markiewicz's BIDS ecosystem presentation

https://bids-specification.readthedocs.io/en/stable/
https://bids.neuroimaging.io/
https://bids-standard.github.io/bids-validator/

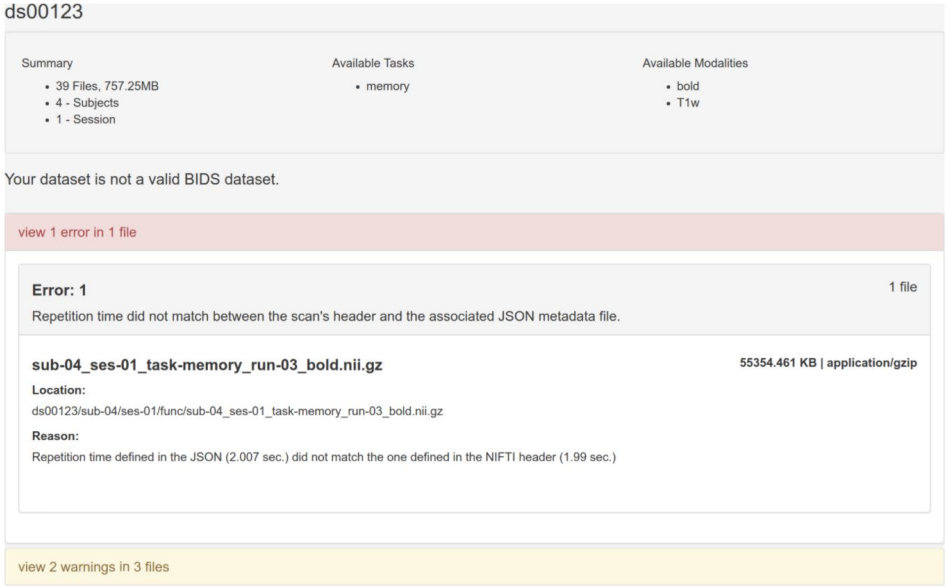
- BIDS' machine readable structure makes automated validation possible and easy
- BIDS-validator does a comprehensive check
- for every single file and aspect in dataset:
  - warnings
  - errors
- fast and reliable integrity check

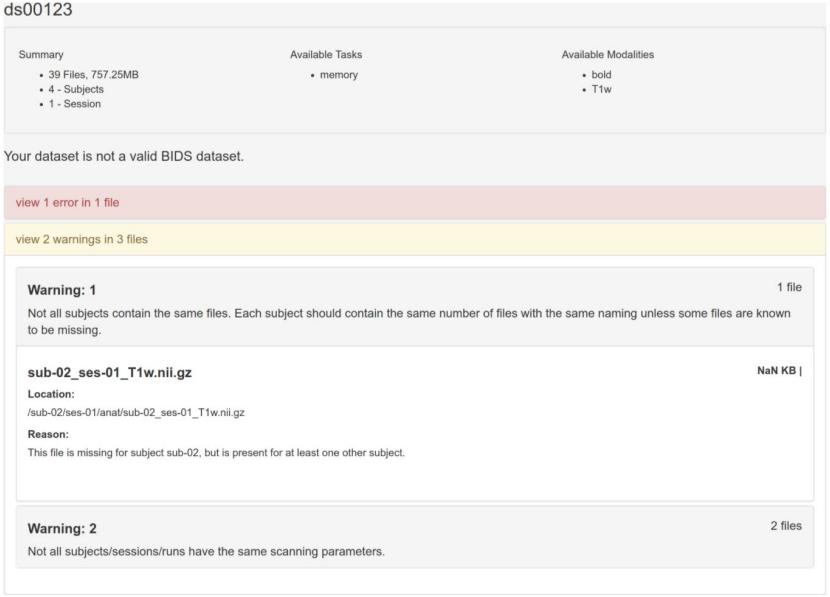


FAIR necessities - standardization - the benefits of BIDS - data validation

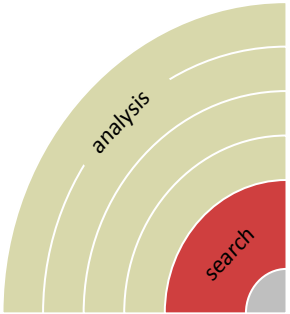
* content liberally borrowed from
Chris Markiewicz's BIDS ecosystem presentation

https://bids-specification.readthedocs.io/en/stable/
https://bids.neuroimaging.io/
https://bids-standard.github.io/bids-validator/





FAIR data through community-driven development of standards and beyond



- necessities - documentation
- **necessities - virtualization (standardized)**
- **necessities - workflows (standardized)**
- necessities - validation (standardized)
- necessities - standardization
- prerequisites - data sharing (standardized)
- prerequisites - preregistration (standardized)

BIDS Apps

Neuroscout

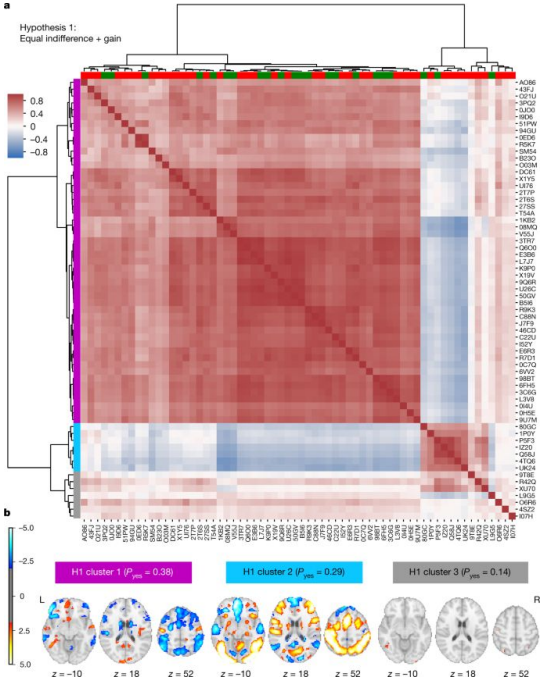
BIDS BRAIN IMAGING DATA STRUCTURE

DOI: 10.5281/zenodo.1147025

FAIR necessities - standardization - the benefits of BIDS - data processing

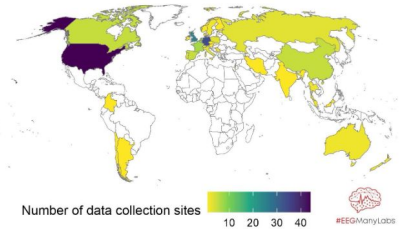
- a lot of sources of analyses/results variability
- every aspect of scientific workflow matters and therefore should be **FAIR**
- acquisition, datasets and metadata
- computing environment (OS, software package, version number, etc.)
- mega & multiverse analyses demonstrated the need for **FAIR** and **standardization**

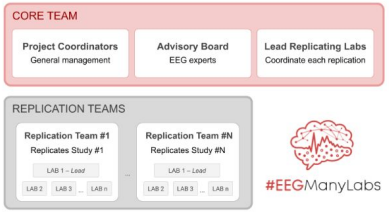
Botvinik-Nezer et al. (2020)
The NARPS project



https://www.nature.com/articles/s41586-020-2314-9

Pavlov et al. (2020)
#EEGManyLabs

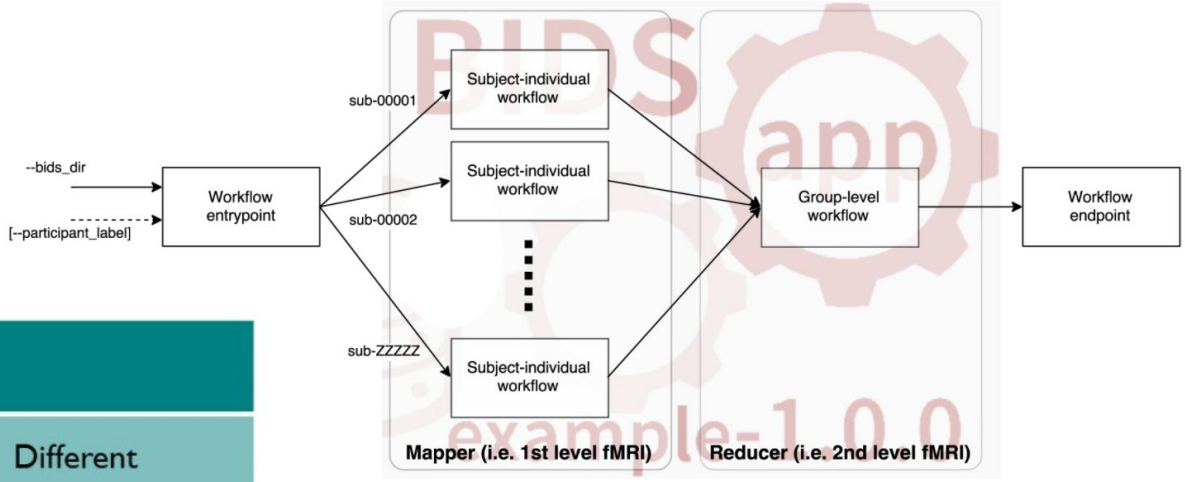




https://psyarxiv.com/528nr/

FAIR necessities - standardization - the benefits of BIDS - data processing

* content liberally borrowed from
  Chris Markiewicz's BIDS ecosystem presentation



From doi:10.1371/journal.pcbi.1005209.g002

Analysis	Data: Same	Data: Different
Same	Reproducible	Replicable
Different	Robust	Generalisable

From The Turing Way, Ch. 2; doi:10.5281/zenodo.3233853

- BIDS Apps aims to address aspects of this problem via a common interface, container technologies and standardization
- open & reproducible analyses on (open) data
- each container/app includes a dedicated pipeline/workflow
- input: BIDS dataset
- pipeline/workflow: virtualization using containers
- output: BIDS conform standardized derivatives
- standardization + flexibility + reproducibility

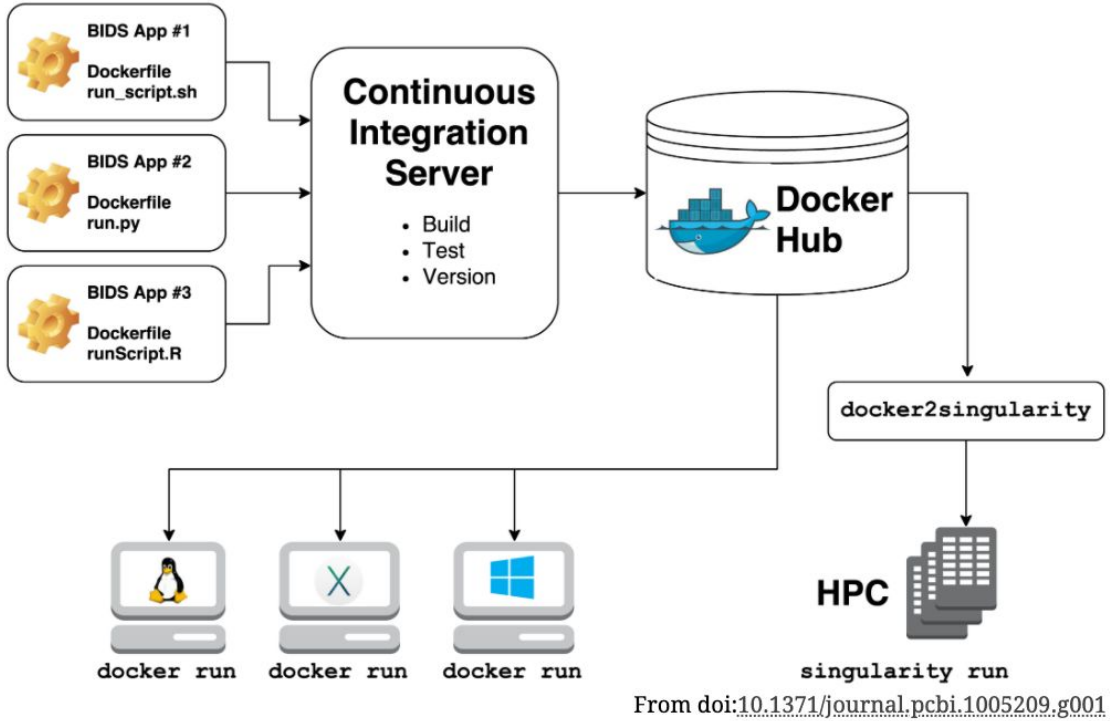
https://bids-apps.neuroimaging.io/

FAIR necessities - standardization - the benefits of BIDS - data processing

* content liberally borrowed from
Chris Markiewicz's BIDS ecosystem presentation

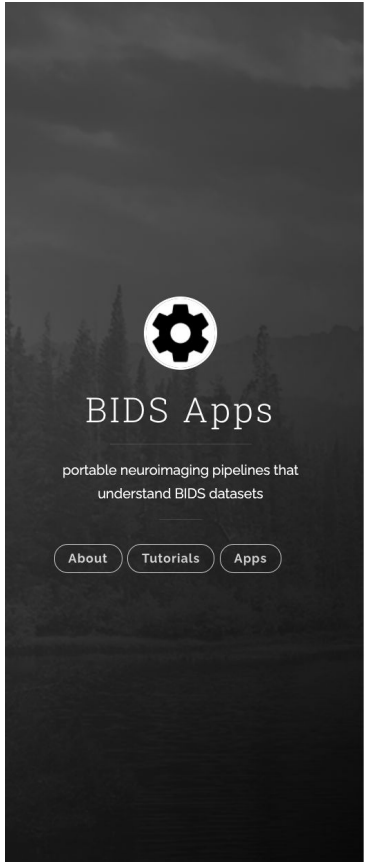
https://bids-apps.neuroimaging.io/

- containerized pipelines/workflow that understand BIDS datasets
- set of core command line arguments
- should not depend on software outside the container
- openly hosted on docker/singularity hub and include version tags
- targeting automated and standardized analyses that are reproducible



From doi:10.1371/journal.pcbi.1005209.g001

FAIR necessities - standardization - the benefits of BIDS - data processing



Available BIDS Apps

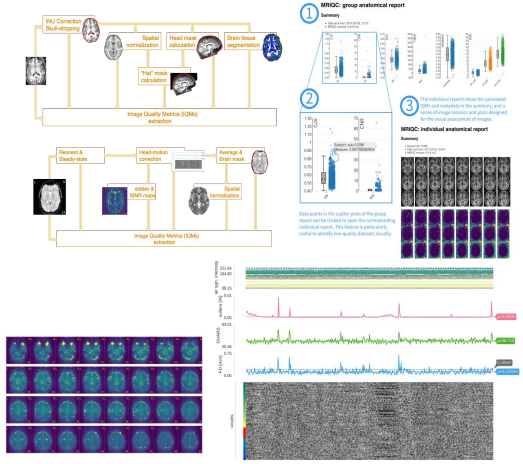
https://bids-apps.neuroimaging.io/

BIDS-Apps/example	version v0.0.7	open issues 5	circleci passing	open bug pull requests 0	docker pulls 14k	439.5MB 23 layers
BIDS-Apps/freesurfer	version v6.0.1-6.1	open issues 13	circleci passing	open bug pull requests 0	docker pulls 11k	2.6GB 52 layers
BIDS-Apps/ndmg	version v0.1.0	open issues 1	circleci passing	open bug pull requests 0	docker pulls 8k	920.9MB 31 layers
BIDS-Apps/BROCCOLI	version v1.0.1	open issues 4	circleci passing	open bug pull requests 0	docker pulls 522	3GB 21 layers
BIDS-Apps/FibreDensityAndCrosssection	version v0.0.1	open issues 1	circleci passing	open bug pull requests 0	docker pulls 250	576.8MB 31 layers
BIDS-Apps/SPM	version v0.0.20	open issues 3	circleci passing	open bug pull requests 0	docker pulls 1.5k	2GB 24 layers
poldracklab/mriqc	version v0.15.3	open issues 125	circleci passing	open bug pull requests 0	docker pulls 69k	3GB 41 layers
BIDS-Apps/QAP	version v0.0.1	open issues 5	circleci passing	open bug pull requests 0	docker pulls 7	Image not found
BIDS-Apps/CPAC	version v1.0.1a_22	open issues 10	circleci passing	open bug pull requests 0	docker pulls 3.2k	1.6GB 45 layers
BIDS-Apps/hyperalignment	version v0.0.5	open issues 0	circleci passing	open bug pull requests 0	docker pulls 169	Image not found
BIDS-Apps/mindboggle	version v0.0.4-1	open issues 8	circleci passing	open bug pull requests 0	docker pulls 752	1.9GB 81 layers
BIDS-Apps/MRtrix3_connectome	version v0.4.2	open issues 34	circleci passing	open bug pull requests 0	docker pulls 1.4k	7.3GB 56 layers
BIDS-Apps/rs_signal_extract	version v0.1	open issues 0	circleci passing	open bug pull requests 0	docker pulls 238	240MB 17 layers
BIDS-Apps/aa	version v0.2.0	open issues 2	circleci passing	open bug pull requests 0	docker pulls 272	6.3GB 30 layers
BIDS-Apps/niak	version v1.0	open issues 3	circleci passing	open bug pull requests 0	docker pulls 279	2.7GB 103 layers
BIDS-Apps/oppni	version v0.7.0-1	open issues 5	circleci passing	open bug pull requests 0	docker pulls 315	2.9GB 41 layers
poldracklab/fmriprep	version v20.2.0	open issues 192	circleci passing	open bug pull requests 0	docker pulls 376k	5GB 49 layers
BIDS-Apps/brainiak-srm	version initial	open issues 0	circleci failing	open bug pull requests 0	docker pulls 207	559.3MB 13 layers
BIDS-Apps/nipypelines	version v0.3.0	open issues 0	circleci passing	open bug pull requests 0	docker pulls 381	478.1MB 20 layers
BIDS-Apps/HCPPipelines	version v4.1.3	open issues 7	circleci passing	open bug pull requests 0	docker pulls 1.9k	4GB 31 layers
BIDS-Apps/MAGeTbrain	version v0.3.1	open issues 2	circleci passing	open bug pull requests 0	docker pulls 466	Image not found
BIDS-Apps/tracula	version v6.0.0-5	open issues 0	circleci passing	open bug pull requests 0	docker pulls 689	3.4GB 57 layers
BIDS-Apps/baracus	version v1.1.4	open issues 2	circleci passing	open bug pull requests 0	docker pulls 1.2k	Image not found
BIDS-Apps/antsCorticalThickness	version v2.2.0-1	open issues 0	circleci passing	open bug pull requests 0	docker pulls 207	351.9MB 21 layers

FAIR necessities - standardization - the benefits of BIDS - data processing

quality control (MRIQC)

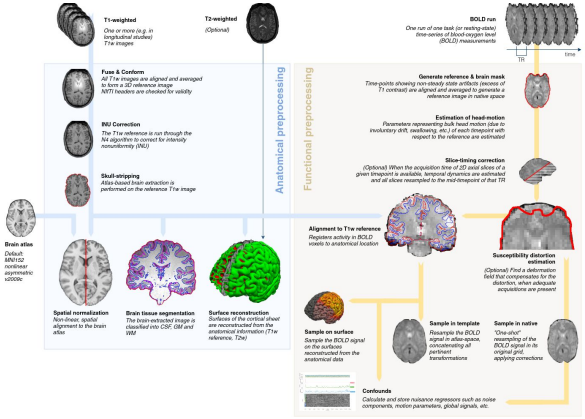
```
mriqc data/ output/ --specifications
```



Esteban et al. (2017),
https://mriqc.readthedocs.io/en/stable/index.html

preprocessing (fMRIPREP)

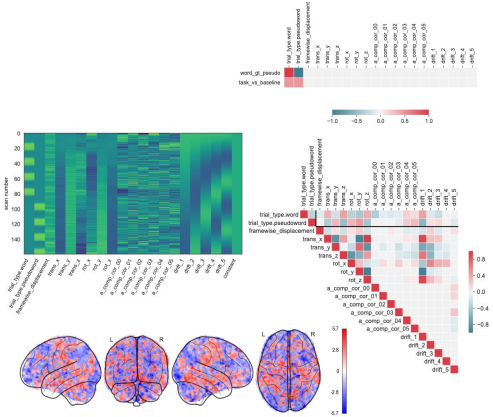
```
fmriprep data/ output/ --specifications
```



Esteban et al. (2019),
https://fmriprep.org/en/stable/

statistical analyses (fitlins)

```
fitlins data/ output/ --specifications
```



https://fitlins.readthedocs.io/en/latest/index.html

FAIR necessities - standardization - the idea of data processing

https://giraffe.tools/

opens up your code

native version control

Doge says: "so much wow:"

clear annotated pipelines

collaborate on your data analysis

show your workflow to your grandparents!

Porcupine

PORcupine Creates Ur PipelINE

With the Porcupine workflow editor, you can visually build your pipeline!

Get started!

Armadillo

Augmented Reality [madillo]

The ARmadillo web app creates 3D Augmented Reality images from the Neurovault database

Get started!

FAIR necessities - standardization - combining data, processing, outcomes

https://neuroscout.org/

Neuroscout

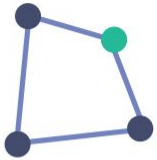
A platform for fast and flexible re-analysis of (naturalistic) fMRI studies

Browse public analyses

Re-use public data

Select from openly available naturalistic fMRI datasets, from sources such as OpenNeuro and DataLad.

Design your analysis



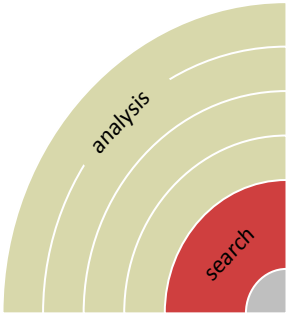
Browse hundreds of annotations automatically extracted from stimuli using **state-of-the-art machine learning** algorithms, such as Google Cloud Vision, IBM Watson, and more.

Execute and share

Portable BIDS pipelines enable execution with no configuration. Results are automatically uploaded to NeuroVault for easy sharing.

Learn more

FAIR data through community-driven development of standards and beyond



- **necessities - documentation**
- necessities - virtualization (standardized)
- necessities - workflows (standardized)

- necessities - standardization
- prerequisites - data sharing (standardized)
- prerequisites - preregistration (standardized)

DOI: 10.5281/zenodo.1147025

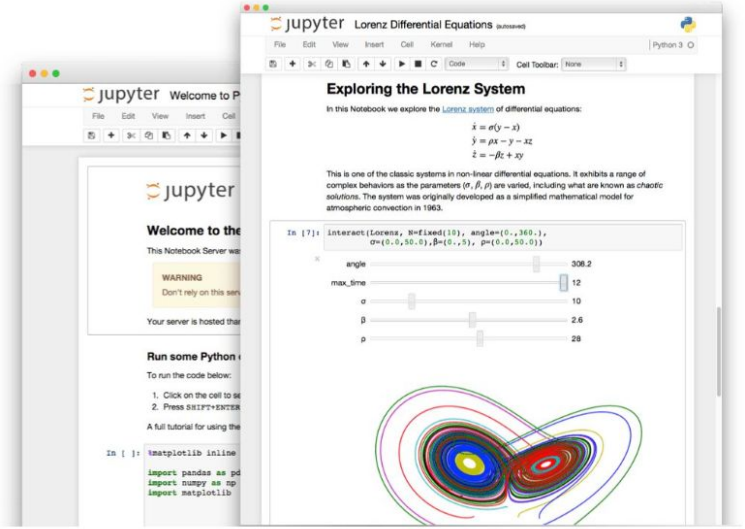
FAIR necessities - analyses documentation - Jupyter

*adapted from Fernando Perez
and Elizabeth DuPre

- a **community** of people
- an **ecosystem** of **open tools** and **standards** for **interactive computing**
- **language-agnostic** and **modular**
- **empower people** to use other **open tools**

FAIR necessities - analyses documentation - Jupyter notebooks

*adapted from Fernando Perez
and Elizabeth DuPre

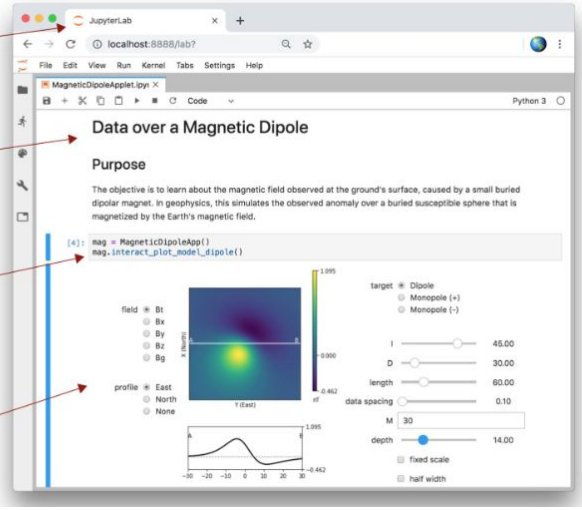


Runs in web browser

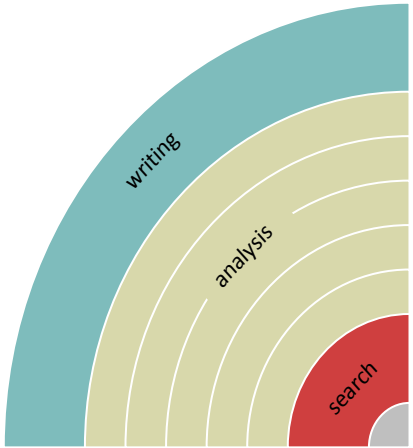
Text, comments, equations

Lines of code
(Python, Julia, R, Matlab, …)

Output: values, images, plots,
tables, interactive widgets



FAIR data through community-driven development of standards and beyond



- **necessities - reporting (standardized)**

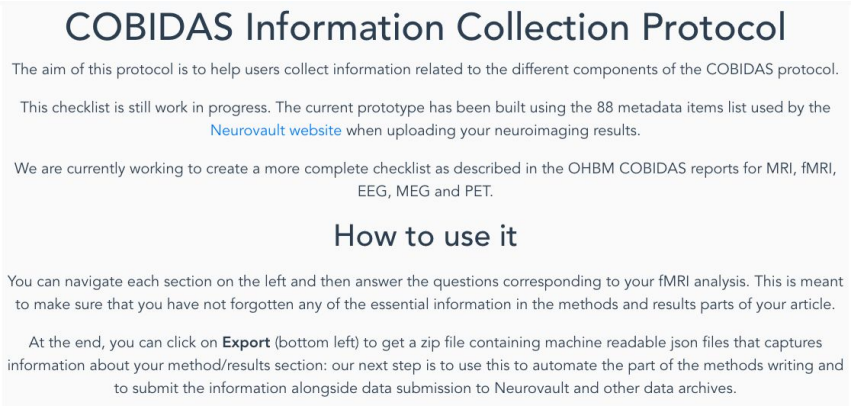
- necessities - documentation
- necessities - virtualization (standardized)
- necessities - workflows (standardized)

- necessities - standardization
- prerequisites - data sharing (standardized)
- prerequisites - preregistration (standardized)

BIDS Apps

DOI: 10.5281/zenodo.1147025

FAIR necessities - reporting (standardized)



COBIDAS Information Collection Protocol

The aim of this protocol is to help users collect information related to the different components of the COBIDAS protocol.

This checklist is still work in progress. The current prototype has been built using the 88 metadata items list used by the Neurovault website when uploading your neuroimaging results.

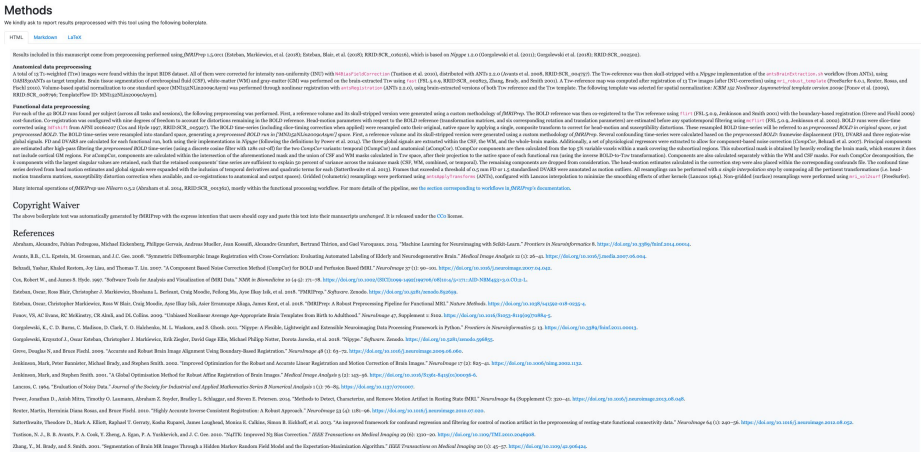
We are currently working to create a more complete checklist as described in the OHBM COBIDAS reports for MRI, fMRI, EEG, MEG and PET.

How to use it

You can navigate each section on the left and then answer the questions corresponding to your fMRI analysis. This is meant to make sure that you have not forgotten any of the essential information in the methods and results parts of your article.

At the end, you can click on **Export** (bottom left) to get a zip file containing machine readable json files that captures information about your method/results section: our next step is to use this to automate the part of the methods writing and to submit the information alongside data submission to Neurovault and other data archives.

https://ohbm.github.io/cobidas/



Methods

We kindly ask to report results preprocessed with this tool using the following boilerplate.

HTML Markdown LaTeX

Results included in this manuscript come from preprocessing performed using *fMRIPrep* 1.5.0rc1 (Esteban, Markiewicz, et al. (2018); Esteban, Blair, et al. (2018); RRID:SCR_016216), which is based on *Nipype* 1.2.0 (Gorgolewski et al. (2011); Gorgolewski et al. (2018); RRID:SCR_002502).

Anatomical data preprocessing

A total of 13 T1-weighted (T1w) images were found within the input BIDS dataset. All of them were corrected for intensity non-uniformity (INU) with `N4BiasFieldCorrection` (Tustison et al. 2010), distributed with ANTs 2.2.0 (Avants et al. 2008, RRID:SCR_004757). The T1w-reference was then skull-stripped with a *Nipype* implementation of the `antsBrainExtraction.sh` workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using `fast` (FSL 5.0.9, RRID:SCR_002823, Zhang, Brady, and Smith 2001). A T1w-reference map was computed after registration of 13 T1w images (after INU-correction) using `mri_robust_template` (FreeSurfer 6.0.1, Reuter, Rosas, and Fischl 2010). Volume-based spatial normalization to one standard space (MNI152NLin2009cAsym) was performed through nonlinear registration with `antsRegistration` (ANTs 2.2.0), using brain-extracted versions of both T1w reference and the T1w template. The following template was selected for spatial normalization: *ICBM 152 Nonlinear Asymmetrical template version 2009c* (Fonov et al. (2009); RRID:SCR_008796; TemplateFlow ID: MNI152NLin2009cAsym).

Functional data preprocessing

For each of the 40 BOLD runs found per subject (across all tasks and sessions), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. The BOLD reference was then co-registered to the T1w reference using `flirt` (FSL 5.0.9, Jenkinson and Smith 2001) with the boundary-based registration (Greve and Fischl 2009) cost-function. Co-registration was configured with nine degrees of freedom to account for distortions remaining in the BOLD reference. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using `mcflirt` (FSL 5.0.9, Jenkinson et al. 2002). BOLD runs were slice-time corrected using `3dTshift` from AFNI 20160207 (Cox and Hyde 1997, RRID:SCR_005927). The BOLD time-series (including slice-timing correction when applied) were resampled onto their original, native space by applying a single, composite transform to correct for head-motion and susceptibility distortions. These resampled BOLD time-series will be referred to as *preprocessed BOLD* in original space, or just *preprocessed BOLD*. The BOLD time-series were resampled into standard space, generating a *preprocessed BOLD run in [MNI152NLin2009cAsym] space*. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. Several confounding time-series were calculated based on the *preprocessed BOLD*: framewise displacement (FD), DVARS and three region-wise global signals. FD and DVARS are calculated for each functional run, both using their implementations in *Nipype* (following the definitions by Power et al. 2014). The three global signals are extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (*CompCor*, Behzadi et al. 2007). Principal components are estimated after high-pass filtering the *preprocessed BOLD* time-series (using a discrete cosine filter with 128s cut-off) for the two *CompCor* variants: temporal (tCompCor) and anatomical (aCompCor). tCompCor components are then calculated from the top 5% variable voxels within a mask covering the subcortical regions. This subcortical mask is obtained by heavily eroding the brain mask, which ensures it does not include cortical GM regions. For aCompCor, components are calculated within the intersection of the aforementioned mask and the union of CSF and WM masks calculated in T1w space, after their projection to the native space of each functional run (using the inverse BOLD-to-T1w transformation). Components are also calculated separately within the WM and CSF masks. For each CompCor decomposition, the k components with the largest singular values are retained, such that the retained components' time series are sufficient to explain 50 percent of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components are dropped from consideration. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each (Satterthwaite et al. 2013). Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardised DVARS were annotated as motion outliers. All resamplings can be performed with a single interpolation step by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using `antsApplyTransforms` (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels (Lanczos 1964). Non-gridded (surface) resamplings were performed using `mri_vol2surf` (FreeSurfer).

Many internal operations of *fMRIPrep* use *Nilearn* 0.5.2 (Abraham et al. 2014, RRID:SCR_001362), mostly within the functional processing workflow. For more details of the pipeline, see the section corresponding to workflows in *fMRIPrep*'s documentation.

Copyright Waiver

The above boilerplate text was automatically generated by fMRIPrep with the express intention that users should copy and paste this text into their manuscripts *unchanged*. It is released under the CC0 license.

References

Abraham, Alexandre, Fabian Pedregosa, Michael Eickenberg, Philippe Gervais, Andreas Mueller, Jean Kossaifi, Alexandre Gramfort, Bertrand Thirion, and Gael Varoquaux. 2014. "Machine Learning for Neuroimaging with Scikit-Learn." *Frontiers in Neuroinformatics* 8. https://doi.org/10.3389/fninf.2014.00014.

Avants, B.B., C.L. Epstein, M. Grossman, and J.C. Gee. 2008. "Symmetric Diffeomorphic Image Registration with Cross-Correlation: Evaluating Automated Labeling of Elderly and Neurodegenerative Brain." *Medical Image Analysis* 12 (1): 26–41. https://doi.org/10.1016/j.media.2007.06.004.

Behzadi, Yashar, Khaled Restom, Joy Liau, and Thomas T. Liu. 2007. "A Component Based Noise Correction Method (CompCor) for BOLD and Perfusion Based fMRI." *NeuroImage* 37 (1): 90–101. https://doi.org/10.1016/j.neuroimage.2007.04.042.

Cox, Robert W., and James S. Hyde. 1997. "Software Tools for Analysis and Visualization of fMRI Data." *NMR in Biomedicine* 10 (4-5): 171–78. https://doi.org/10.1002/(SICI)1099-1492(199706/08)10:4/5<171::AID-NBM468>3.0.CO;2-L.

Esteban, Oscar, Ross Blair, Christopher J. Markiewicz, Shoshana L. Berleant, Craig Moodie, Feilong Ma, Ayse Ilkay Isik, et al. 2018. "FMRIPrep." *Software*. Zenodo. https://doi.org/10.5281/zenodo.852659.

Esteban, Oscar, Christopher Markiewicz, Ross W Blair, Craig Moodie, Ayse Ilkay Isik, Asier Erramuzpe Aliaga, James Kent, et al. 2018. "FMRIPrep: A Robust Preprocessing Pipeline for Functional MRI." *Nature Methods*. https://doi.org/10.1038/s41592-018-0235-4.

Fonov, VS, AC Evans, RC McKinstry, CR Almli, and DL Collins. 2009. "Unbiased Nonlinear Average Age-Appropriate Brain Templates from Birth to Adulthood." *NeuroImage* 47, Supplement 1: S102. https://doi.org/10.1016/S1053-8119(09)70884-5.

Gorgolewski, K., C. D. Burns, C. Madison, D. Clark, Y. O. Halchenko, M. L. Waskom, and S. Ghosh. 2011. "Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python." *Frontiers in Neuroinformatics* 5: 13. https://doi.org/10.3389/fninf.2011.00013.

Gorgolewski, Krzysztof J., Oscar Esteban, Christopher J. Markiewicz, Erik Ziegler, David Gage Ellis, Michael Philipp Notter, Dorota Jarecka, et al. 2018. "Nipype." *Software*. Zenodo. https://doi.org/10.5281/zenodo.596855.

Greve, Douglas N. and Bruce Fischl. 2009. "Accurate and Robust Brain Image Alignment Using Boundary-Based Registration." *NeuroImage* 48 (1): 63–72. https://doi.org/10.1016/j.neuroimage.2009.06.060.

Jenkinson, Mark, Peter Bannister, Michael Brady, and Stephen Smith. 2002. "Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images." *NeuroImage* 17 (2): 825–41. https://doi.org/10.1006/nimg.2002.1132.

Jenkinson, Mark, and Stephen Smith. 2001. "A Global Optimisation Method for Robust Affine Registration of Brain Images." *Medical Image Analysis* 5 (2): 143–56. https://doi.org/10.1016/S1361-8415(01)00036-6.

Lanczos, C. 1964. "Evaluation of Noisy Data." *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis* 1 (1): 76–85. https://doi.org/10.1137/0701007.

Power, Jonathan D., Anish Mitra, Timothy O. Laumann, Abraham Z. Snyder, Bradley L. Schlaggar, and Steven E. Petersen. 2014. "Methods to Detect, Characterize, and Remove Motion Artifact in Resting State fMRI." *NeuroImage* 84 (Supplement C): 320–41. https://doi.org/10.1016/j.neuroimage.2013.08.048.

Reuter, Martin, Hermioni Diana Rosas, and Bruce Fischl. 2010. "Highly Accurate Inverse Consistent Registration: A Robust Approach." *NeuroImage* 53 (4): 1181–96. https://doi.org/10.1016/j.neuroimage.2010.07.020.

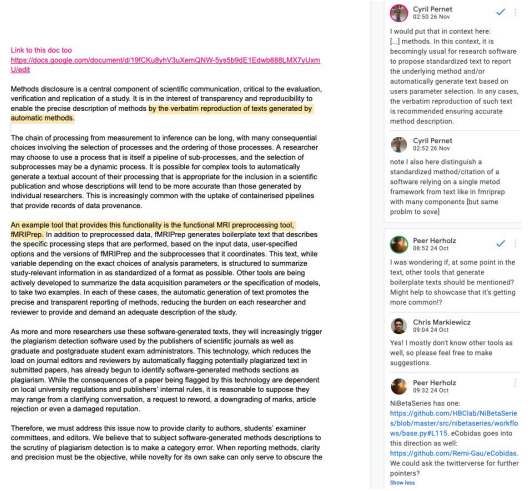
Satterthwaite, Theodore D., Mark A. Elliott, Raphael T. Gerraty, Kosha Ruparel, James Loughead, Monica E. Calkins, Simon B. Eickhoff, et al. 2013. "An Improved Framework for Confound Regression and Filtering for Control of Motion Artifact in the Preprocessing of Resting-State Functional Connectivity Data." *NeuroImage* 64 (1): 240–56. https://doi.org/10.1016/j.neuroimage.2012.08.052.

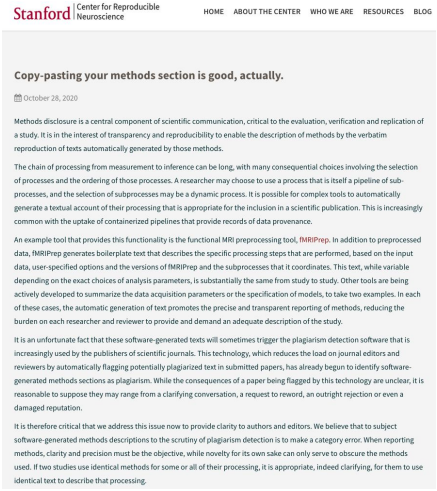
Tustison, N. J., B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee. 2010. "N4ITK: Improved N3 Bias Correction." *IEEE Transactions on Medical Imaging* 29 (6): 1310–20. https://doi.org/10.1109/TMI.2010.2046908.

Zhang, Y., M. Brady, and S. Smith. 2001. "Segmentation of Brain MR Images Through a Hidden Markov Random Field Model and the Expectation-Maximization Algorithm." *IEEE Transactions on Medical Imaging* 20 (1): 45–57. https://doi.org/10.1109/42.906424.

https://fmriprep.org/en/stable/outputs.html

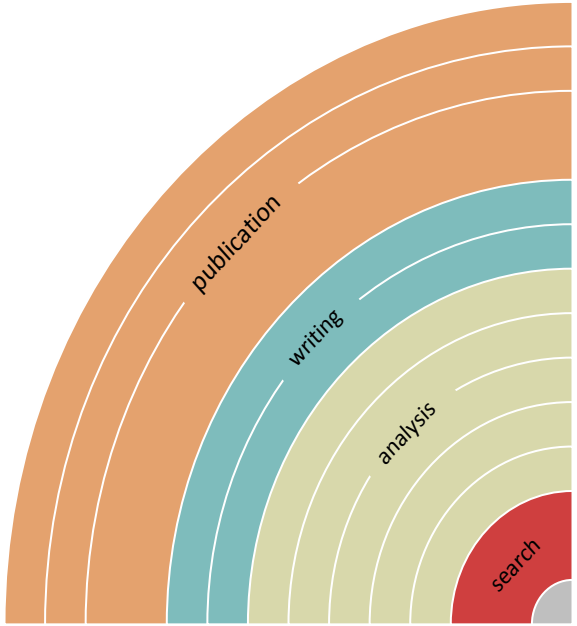
FAIR necessities - reporting (standardized)





- plagiarism software flags automatically generated method sections
- community gets together to gather information and draft statement to address the issue
- statement posted, continuing working together with journals

FAIR data through community-driven development of standards and beyond



- necessities - repositories (standardized)
- **necessities - more than a PDF**

- necessities - reporting (standardized)

- necessities - documentation
- necessities - virtualization (standardized)
- necessities - workflows (standardized)

- necessities - standardization
- prerequisites - data sharing (standardized)
- prerequisites - preregistration (standardized)

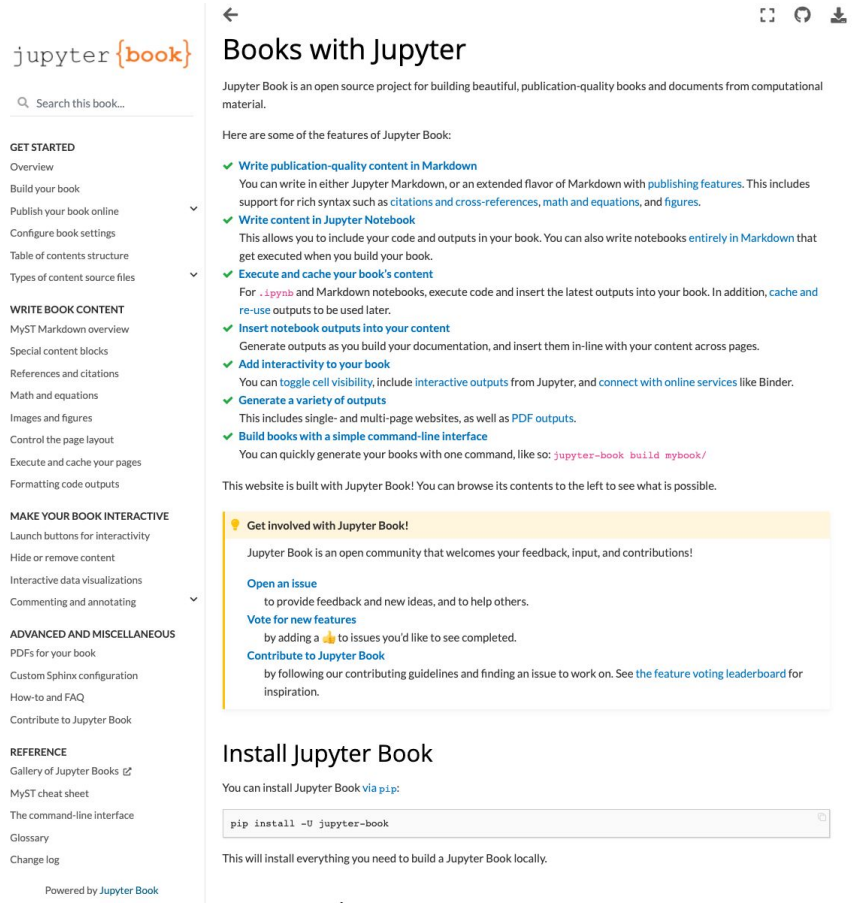
DOI: 10.5281/zenodo.1147025

FAIR necessities - more than a PDF - binder

https://mybinder.org/

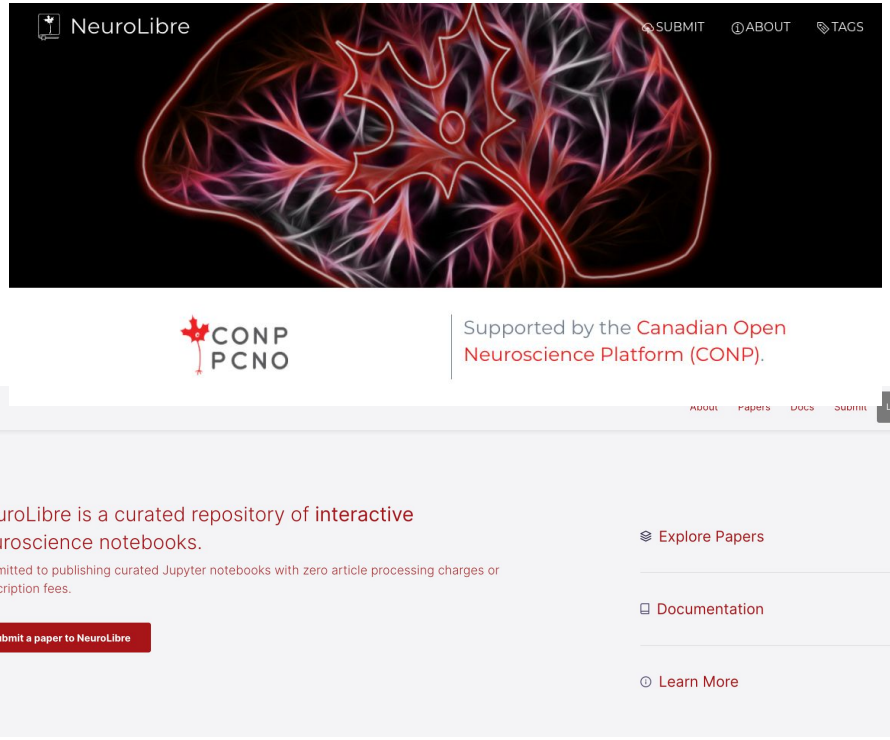
*adapted from Fernando Perez

FAIR necessities - more than a PDF - Jupyter Books



https://jupyterbook.org/

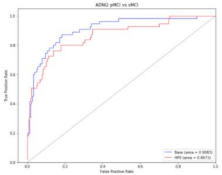
FAIR necessities - more than a PDF - Jupyter + binder



https://www.neurolibre.com/

A highly predictive signature (HPS) of Alzheimer's disease dementia from cognitive and structural brain features

A jupyter notebook containing analyses that give a highly predictive signature (HPS) of Alzheimer's disease dementia from cognitive and structural features using simulated data.



Introduction to machine learning with Nilearn

An introductory tutorial for using the popular Nilearn software package to perform machine learning analyses with neuroimaging data. This material is adapted from the Montreal AI and Neuroscience (MAIN) 2018 workshops.

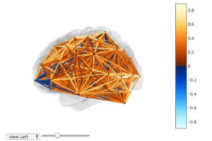
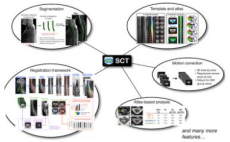
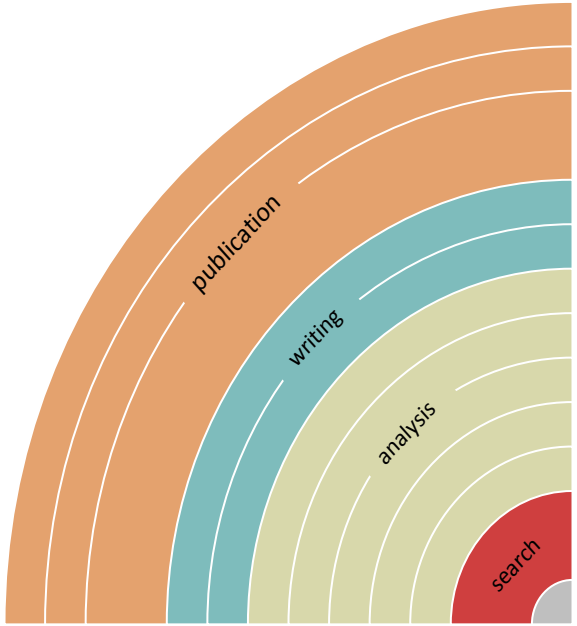


Image processing with Spinal Cord Toolbox (SCT)

This notebook presents an example analysis pipeline using the Spinal Cord Toolbox (SCT), a suite of tools specialized for analysis of spinal cord MRI images of the spinal. Topics covered include: segmentation, masking, registration, warping, and quantitative metric computation. This tutorial was generated in a Jupyter Notebook and coded in Python.



FAIR data through community-driven development of standards and beyond



- **necessities - repositories (standardized)**
- necessities - more than a PDF

- necessities - reporting (standardized)

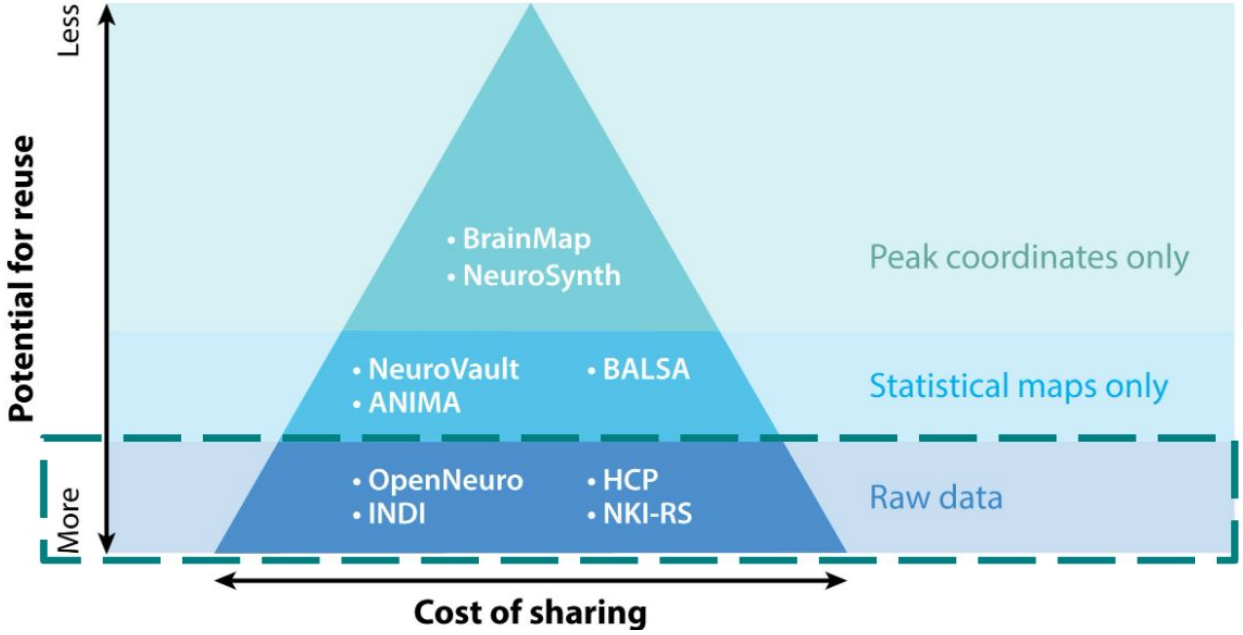
- necessities - documentation
- necessities - virtualization (standardized)
- necessities - workflows (standardized)

- necessities - standardization
- prerequisites - data sharing (standardized)
- prerequisites - preregistration (standardized)

OpenNEURO

DOI: 10.5281/zenodo.1147025

FAIR data through community-driven development of standards and beyond



*adapted from Russ Poldrack

FAIR necessities - repositories (standardized)

OpenNEURO

A free and open platform for sharing MRI, MEG, EEG, iEEG, and ECoG data

Search Datasets

Browse All Public Datasets

436
Public Datasets

14277
Participants

Get Data

Browse and download datasets from contributors all over the world.

Share Data

Upload your data to an NIH Brain Initiative approved repository.

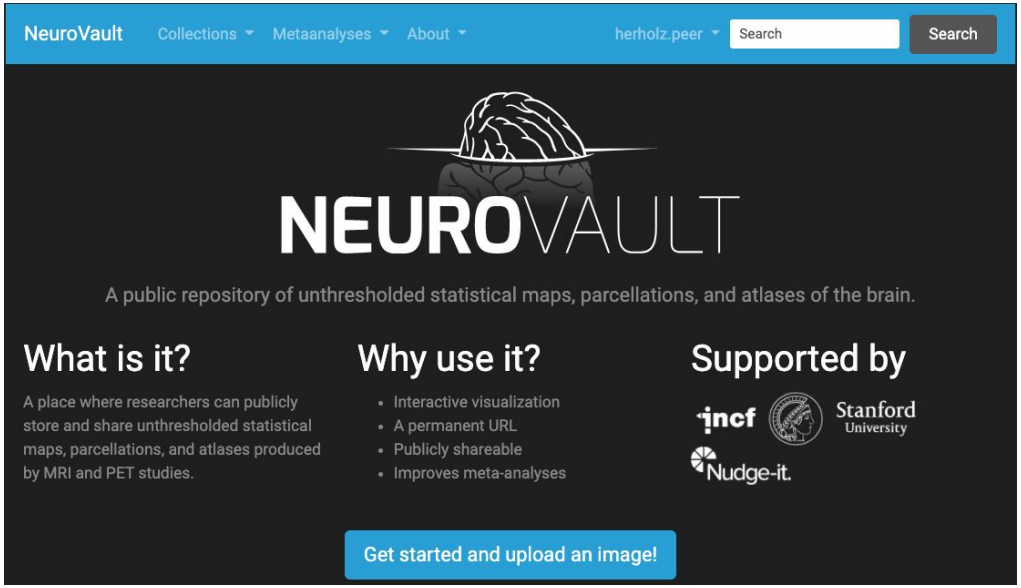
Use Data

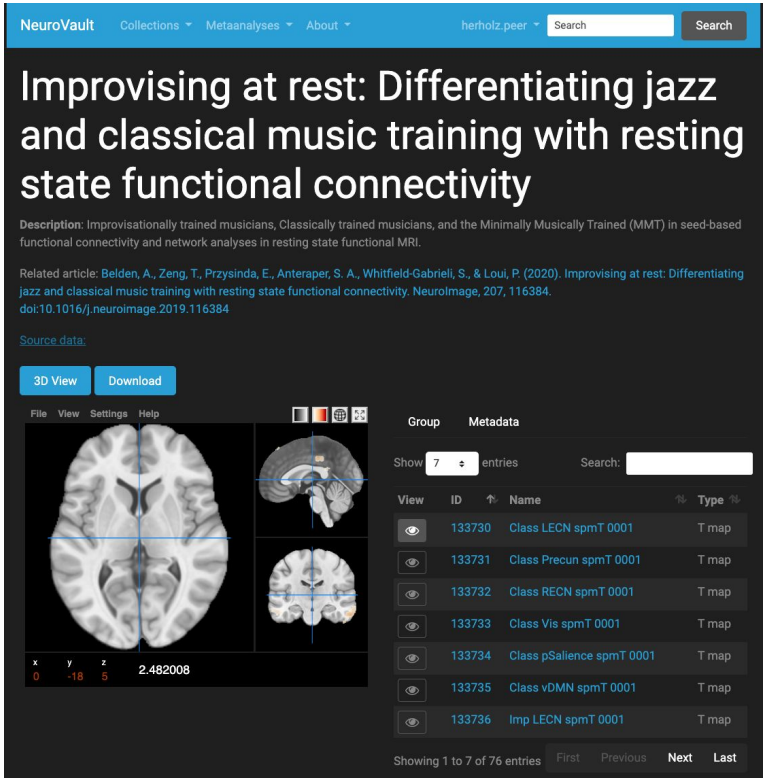
Use our affiliated website to process applicable data.

https://openneuro.org/

FAIR necessities - repositories (standardized)

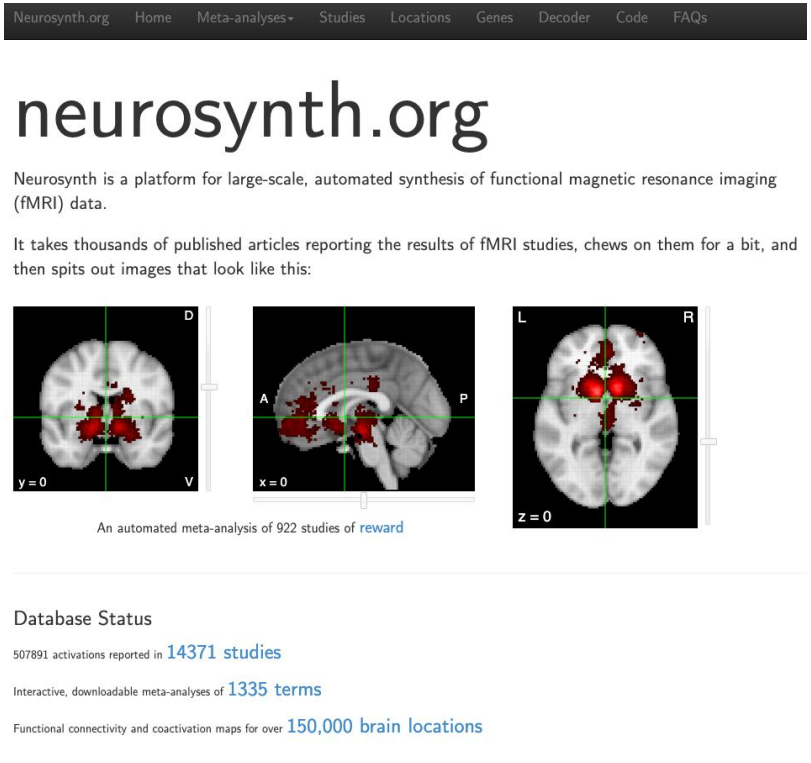
https://neurovault.org/

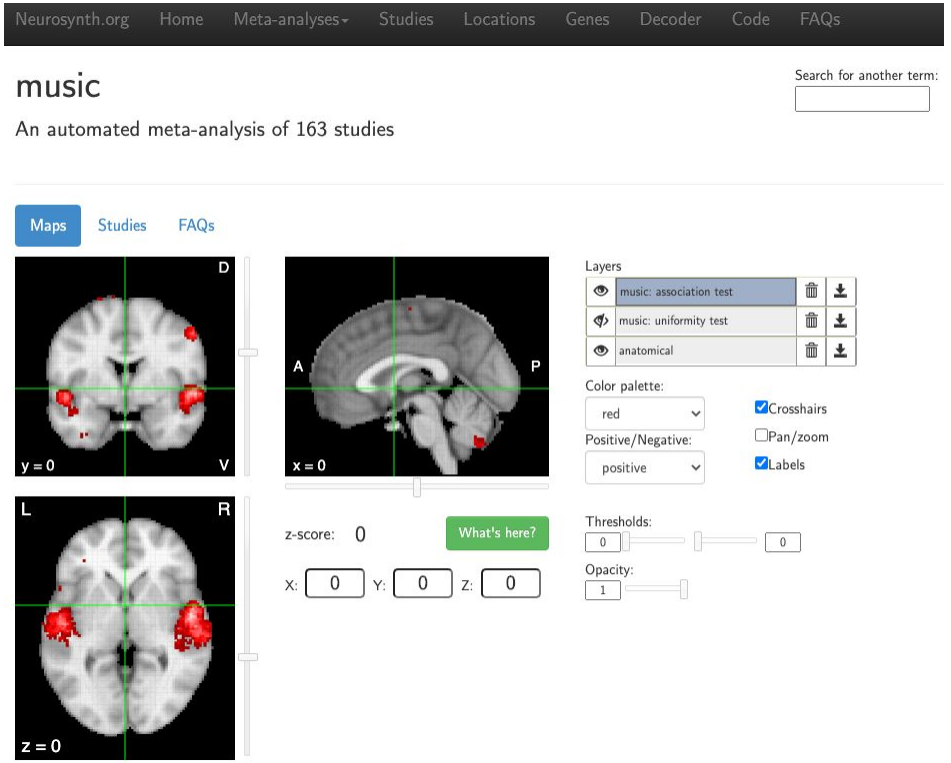




FAIR necessities - repositories (standardized)

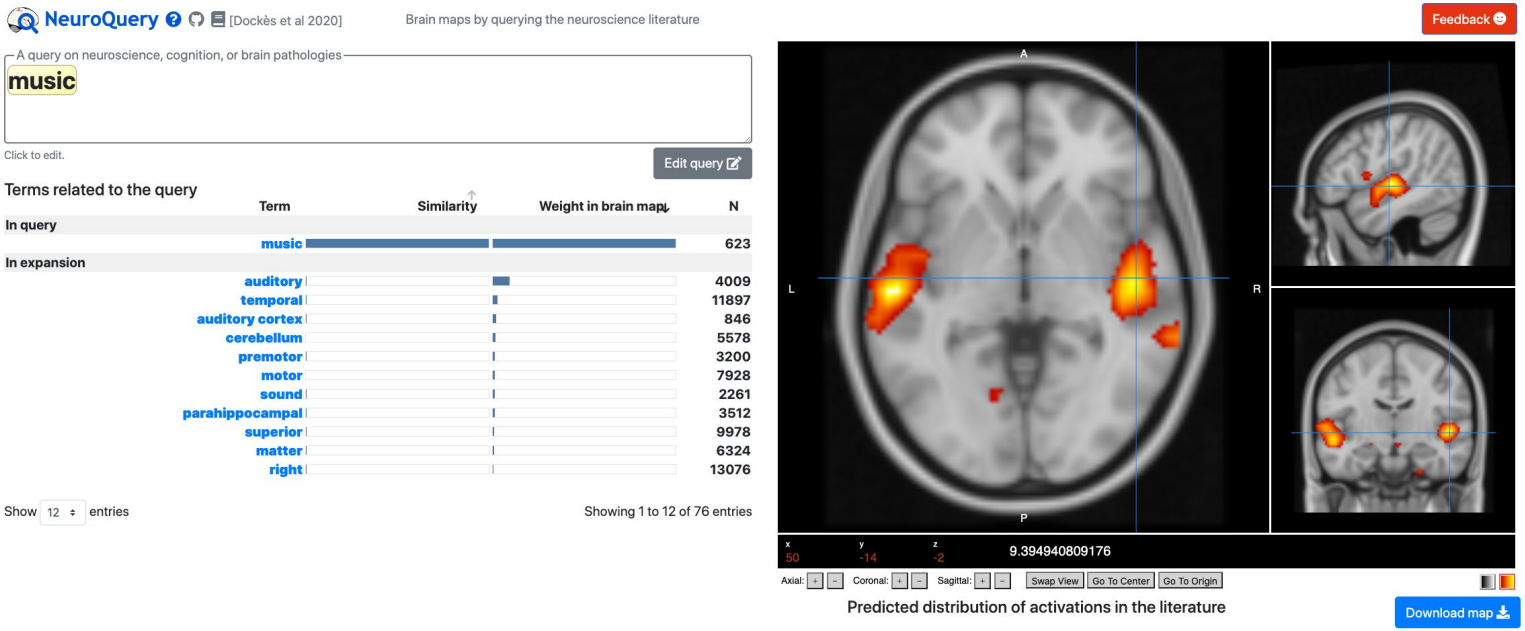
https://www.neurosynth.org/





FAIR necessities - repositories (standardized)

https://neuroquery.org/



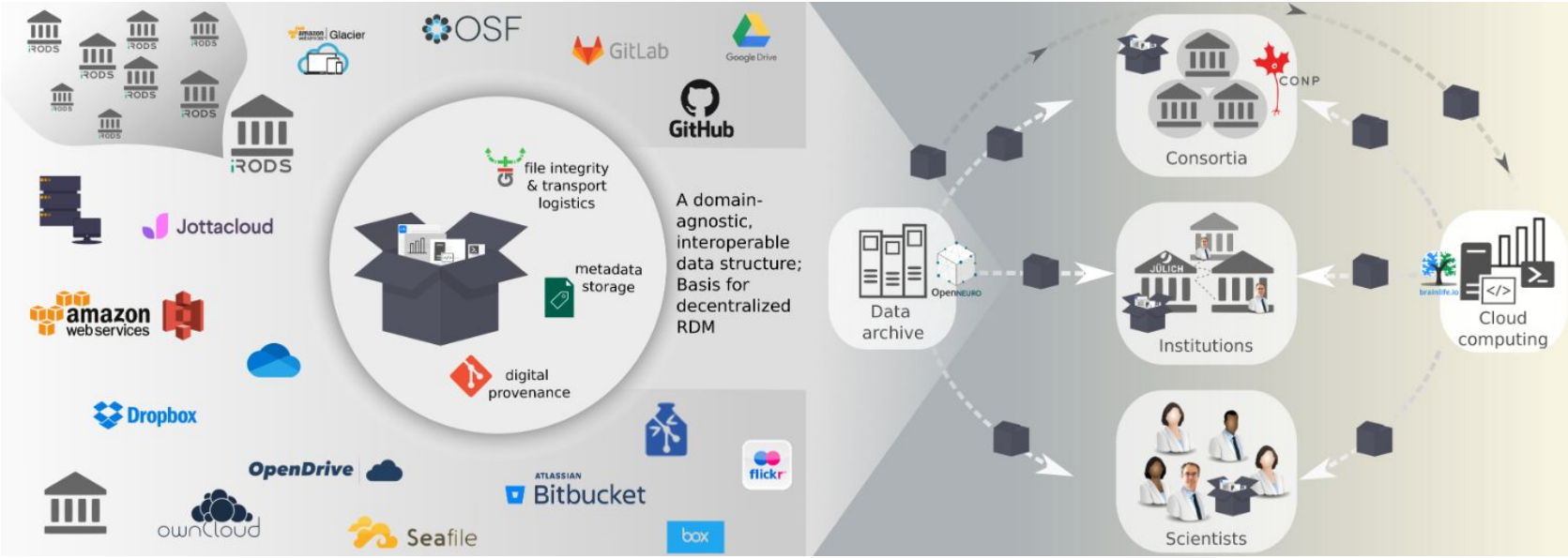
Publications related to the query

- Separate cortical networks involved in music perception: preliminary functional MRI evidence for modularity of music processing
- Music and Emotions in the Brain: Familiarity Matters
- It don't mean a thing...
- A Functional MRI Study of Happy and Sad Emotions in Music with and without Lyrics
- The rewards of music listening: Response and physiological connectivity of the mesolimbic system
- Moving to Music: Effects of Heard and Imagined Musical Cues on Movement-Related Brain Activity

FAIR data management

*adapted from Adina Wagner

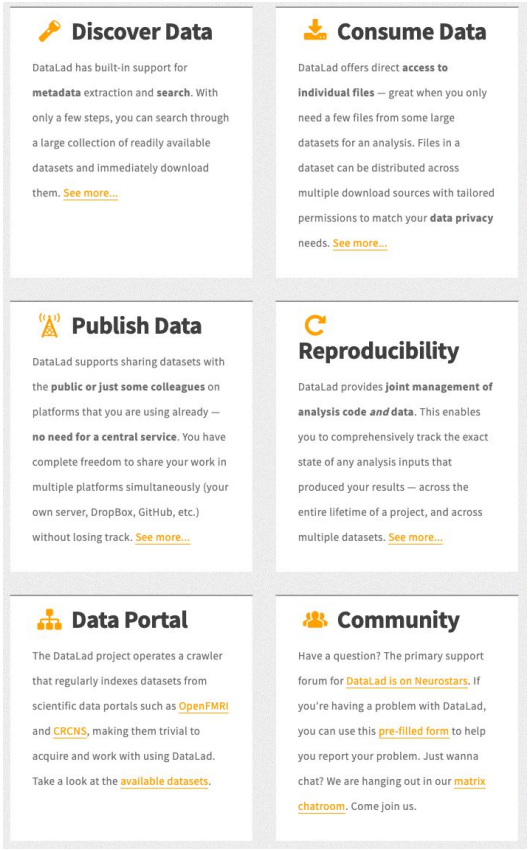
https://www.datalad.org
http://handbook.datalad.org

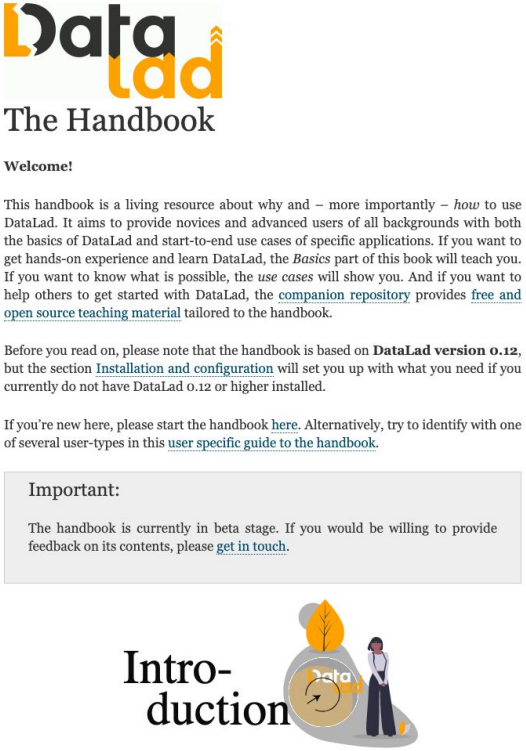


FAIR data management

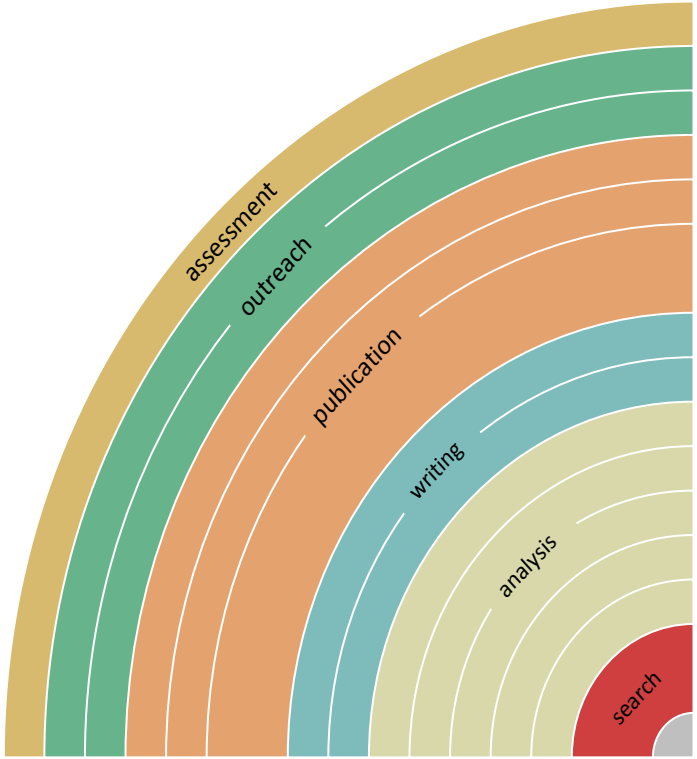
https://www.datalad.org

http://handbook.datalad.org





FAIR data through community-driven development of standards and beyond



- **necessities - support**
- **prerequisites - initiatives**

- necessities - repositories (standardized)
- necessities - more than a PDF

- necessities - reporting (standardized)

- necessities - documentation
- necessities - virtualization (standardized)
- necessities - workflows (standardized)
- necessities - validation (standardized)
- necessities - standardization
- prerequisites - data sharing (standardized)
- prerequisites - preregistration (standardized)

DOI: 10.5281/zenodo.1147025

FAIR initiatives & support

Project T(eaching) **I**(ntegrity in) **E**(mpircal) **R**(eseach)

Raw Data

A copy of **every original data file** from which you extract any of data used in your study.

Your original data files **serve as a record** of the data you began the project with.

Analysis Code + Analysis Data

One or more files containing code used for the study... should **execute all the data processing and analysis necessary to replicate the study** and reproduce the reported results

Documents

Documentation to understand the study

- A copy of your final paper
- Your Data Appendix
- Your README file

https://www.projecttier.org/

The Turing Way

https://zenodo.org/record/3695300#.X8qi_6pKiUI

https://the-turing-way.netlify.app/

FAIR initiatives & support

https://www.repronim.org

ReproNim DISCOVER DESCRIBE DO TEACH COMMUNITY

ReproNim: A Center for Reproducible Neuroimaging Computation

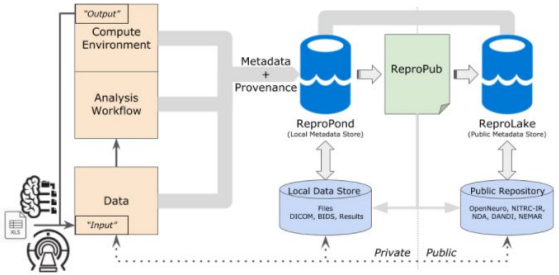
Welcome to ReproNim!

ReproNim's goal is to improve the reproducibility of neuroimaging science and extend the value of our national investment in neuroimaging research, while making the process easier and more efficient for investigators.

ReproNim delivers a reproducible analysis framework comprised of components that include:

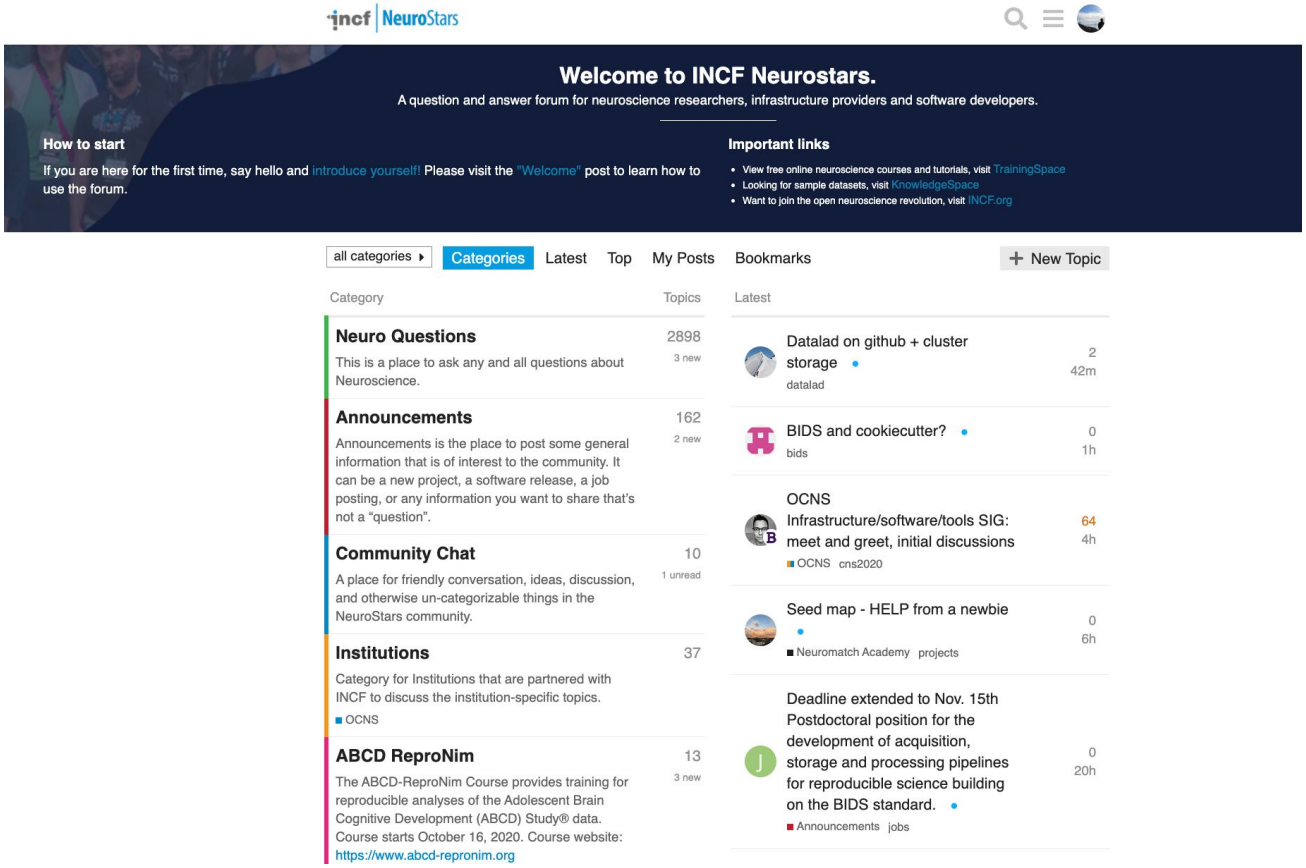
- data and software discovery
- implementation of standardized description of data, results and workflows
- development of execution options that facilitates operation in all computational environments
- provide training and education to the community

All components of the framework are intended to foster continued use and development of the reproducible and generalizable framework in neuroimaging research. See our Blog Post for more discussion of ReproLake and ReproPond.



FAIR initiatives & support

https://neurostars.org/

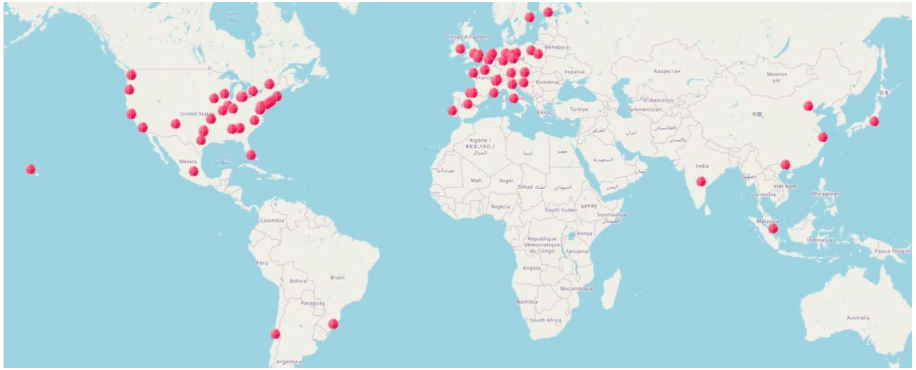


McGill University neuro Brams International Laboratory for Brain, Music, and Sound Research

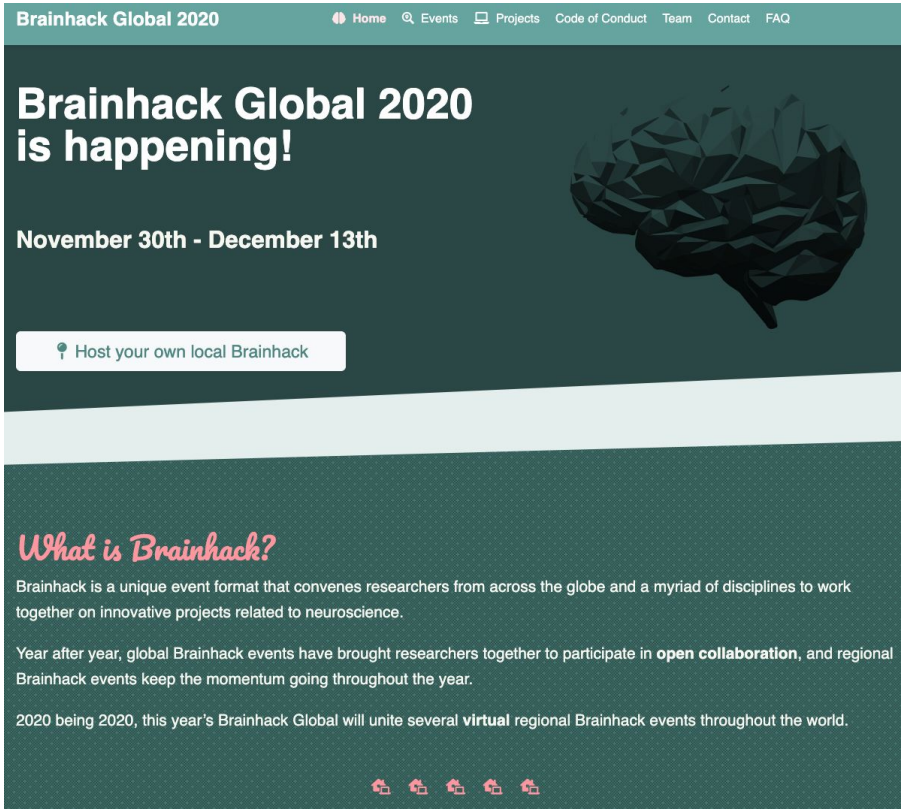
FAIR initiatives & support

https://brainhack.org/global2020/

http://brainhack.org/



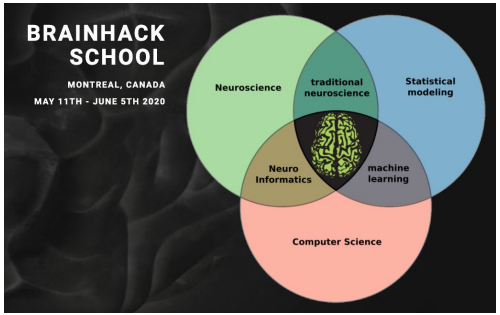
- started in 2012, spread across 25 countries, over 140 events, over 4000 members
- over 1000 attendees annually, focus on ECRs-collaboration-community-hacking-teaching
- model for various large scale teaching initiatives



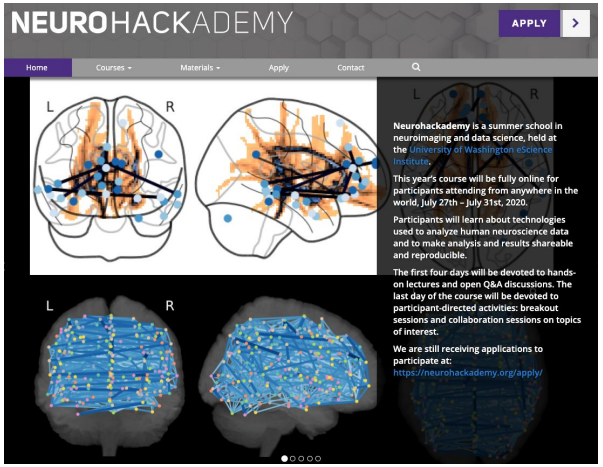
McGill University neuro Brams International Laboratory for Brain, Music, and Sound Research

FAIR initiatives & support

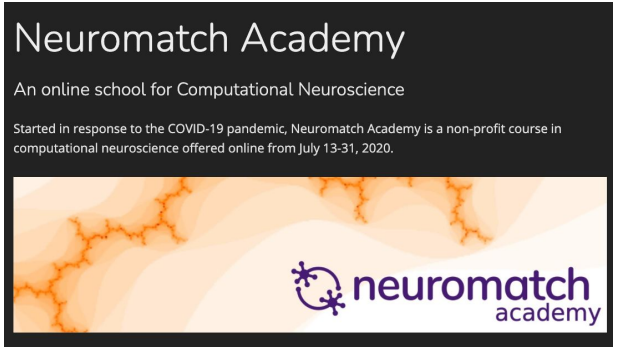
- **open** to everyone interested
- mostly **free** (if not, dedicated funding to support individuals from underrepresented and marginalized groups exists)
- **all materials publicly available** (including video recordings of lectures)
- **comprehensive, application oriented training**
- **decentralized and digital**
- **individual support and supervision**



https://school.brainhackmtl.org/



https://neurohackademy.org/



https://www.neuromatchacademy.org/

FAIR initiatives & support

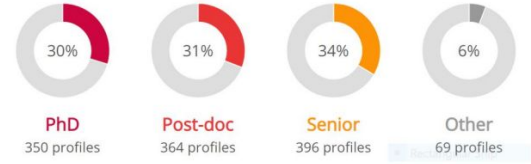
Women in Neuroscience

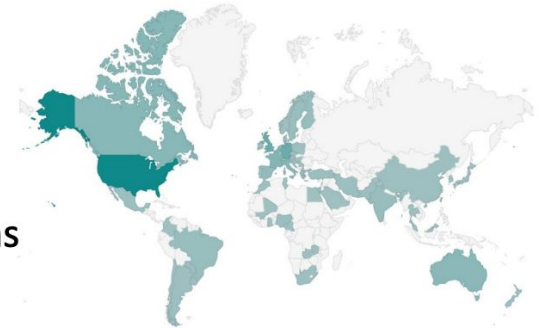
Repository for Women in Neuroscience

- www.winrepo.org
- over 1,100 profiles
- easy search
- recommendations

Support the project:

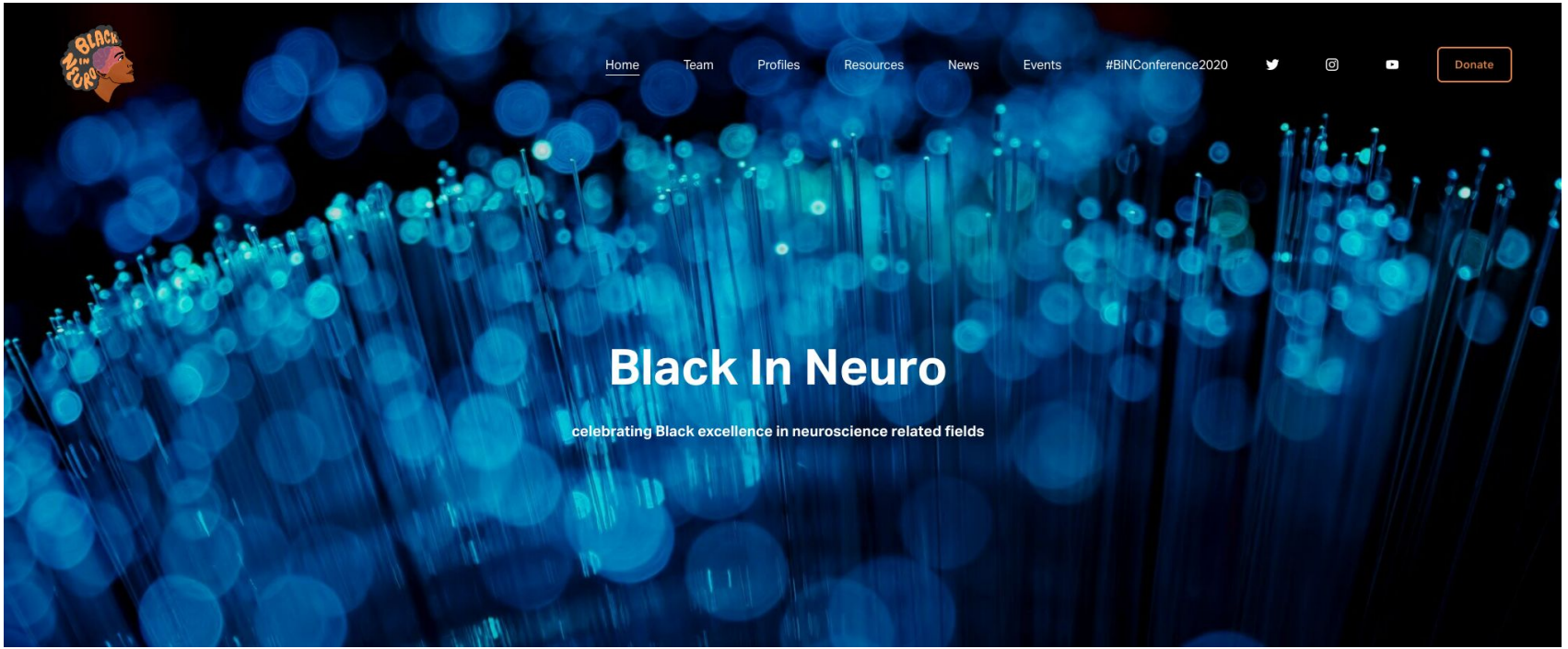
- ➢sign up
- ➢spread the word
- ➢submit recommendations





WINRePo1
www.facebook.com/WiNRepository/

amazing things - super important

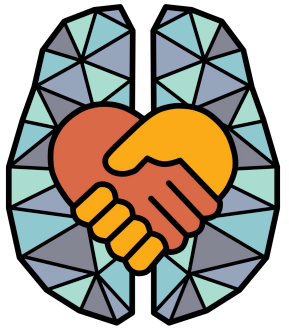


https://www.blackinneuro.com/

FAIR initiatives & support

Queer in Neuro

Google form for signing up



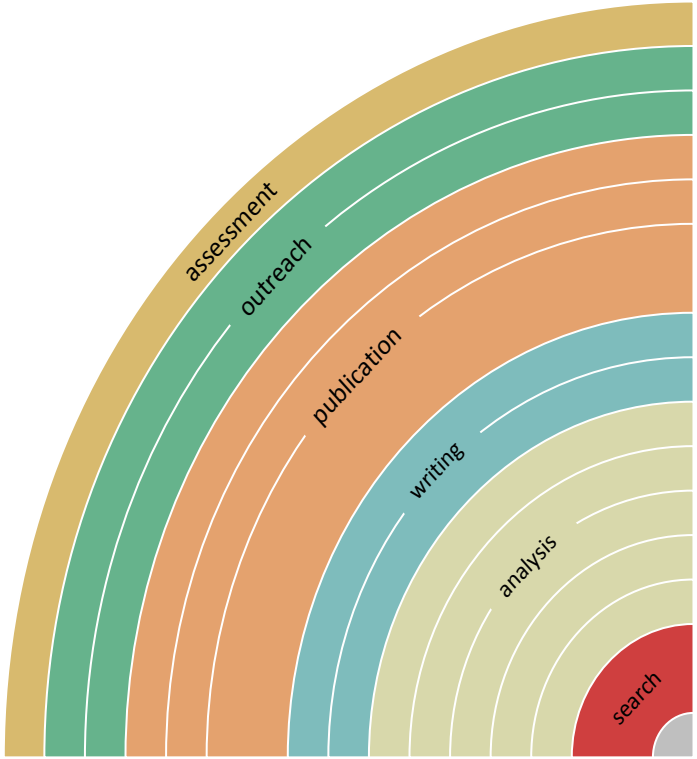
OHBM Student & Postdoc SIG

https://www.ohbmtrainees.com/mentoring-programme/

OSSIG

https://ossig.netlify.app/

FAIR data through community-driven development of standards and beyond



- necessities - support
- prerequisites - initiatives

- necessities - repositories (standardized)
- necessities - more than a PDF

- necessities - reporting (standardized)

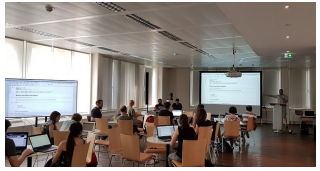
- necessities - documentation
- necessities - virtualization (standardized)
- necessities - workflows (standardized)
- necessities - validation (standardized)
- necessities - standardization
- prerequisites - data sharing (standardized)
- prerequisites - preregistration (standardized)

DOI: 10.5281/zenodo.1147025

FAIR data through community-driven development of standards and beyond

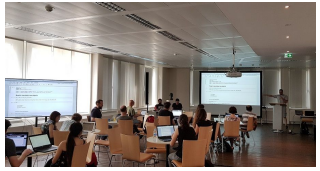
- **community-driven FAIR standards** enable new kinds of science:
  - open and collaborative development that is fast-paced, yet highly structured, understandable and traceable
- **community-driven FAIR standards** introduce changes on a global scale:
  - highly adapted across all levels of individuals and projects
- **community-driven FAIR standards** empower individuals:
  - support and engagement for as many folks with as many backgrounds as possible






FAIR data through community-driven development of standards and beyond

Standardization driven by an **open and structured community** helps **continuous integration**, **development**, **testing**, **adaption**, **extension**, **control**, **broadens science**, **includes everyone**, **teaching/training sustainability**, **getting rid of toxic structures and science islands** and if done early in workflow can **drastically benefit entire research workflow** as every other step can build upon standardization and can itself be standardized.



FAIR

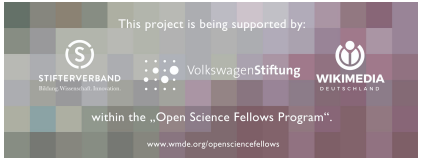
(literally and figuratively)

Big kudos and thank you go out to ...

HEALTHY BRAINS
HEALTHY LIVES

NIH-NIBIB P41 EB019936

National Institute
of Mental Health
R01MH096906



funding

community

Thank you very much for your attention.

I’m happy to answer questions during the chat and/or via twitter (@peerherholz) or email (herholz dot peer at gmail dot com).