

# Open Science with FAIR data and the Brain Imaging Data Structure (BIDS)

Robert Oostenveld

*Donders Institute, Radboud University, Nijmegen, NL*

*NatMEG, Karolinska Institute, Stockholm, SE*



# Open Science

Open access publications

Open peer review

Open educational resources

Open methodology

Open source

Open hardware

Open data



WIKIPEDIA  
The Free Encyclopedia



SCHOLARPEDIA  
the peer-reviewed  
open-access encyclopedia



KHAN  
ACADEMY

edX



PeerJ



biobank<sup>uk</sup>  
Imaging study



HUMAN  
Connectome  
PROJECT

# Making science open = making science better

More efficient and impact – beneficial for you

Higher quality – beneficial for everyone

Pre-registration

Data management plans (think ahead)

Using shared data, sharing your own data

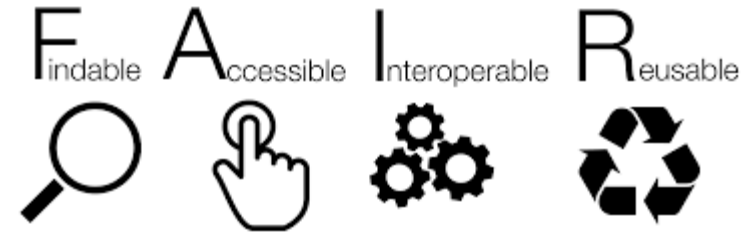
Computational reproducibility

Conceptual replicability

Disclosing all details (code and data)

Open Access publications

# Open Data



## Findable

*Make your data available on repository with a persistent identifier (DOI, handle) and metadata*

## Accessible

*Be explicit about data usage terms (agreement with downloader)*

## Interoperable

*Make your data human and machine readable, e.g. BIDS*

## Reusable

*Make sure you document enough details, e.g. “data descriptor” paper  
this can be cited, along with citing our data -> measurable impact!*



www.nature.com/scientificdata

SCIENTIFIC DATA

OPEN

SUBJECT CATEGORIES

- » Data publication
- » Research article
- » Research data

Received: 18 December 2018  
Accepted: 19 May 2019  
Published: 21 June 2019

Received: 14 November 2018  
Accepted: 3 May 2019  
Published: 19 June 2019

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

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

COMMENT

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

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

# What is it?



BIDS is a way to organize your existing raw data

To improve consistent and complete documentation

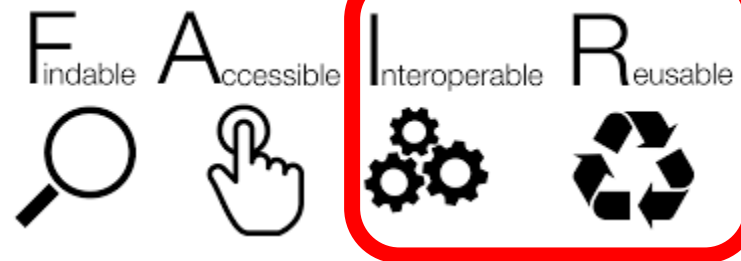
To facilitate re-use by your future self and others

BIDS is not

A new file format

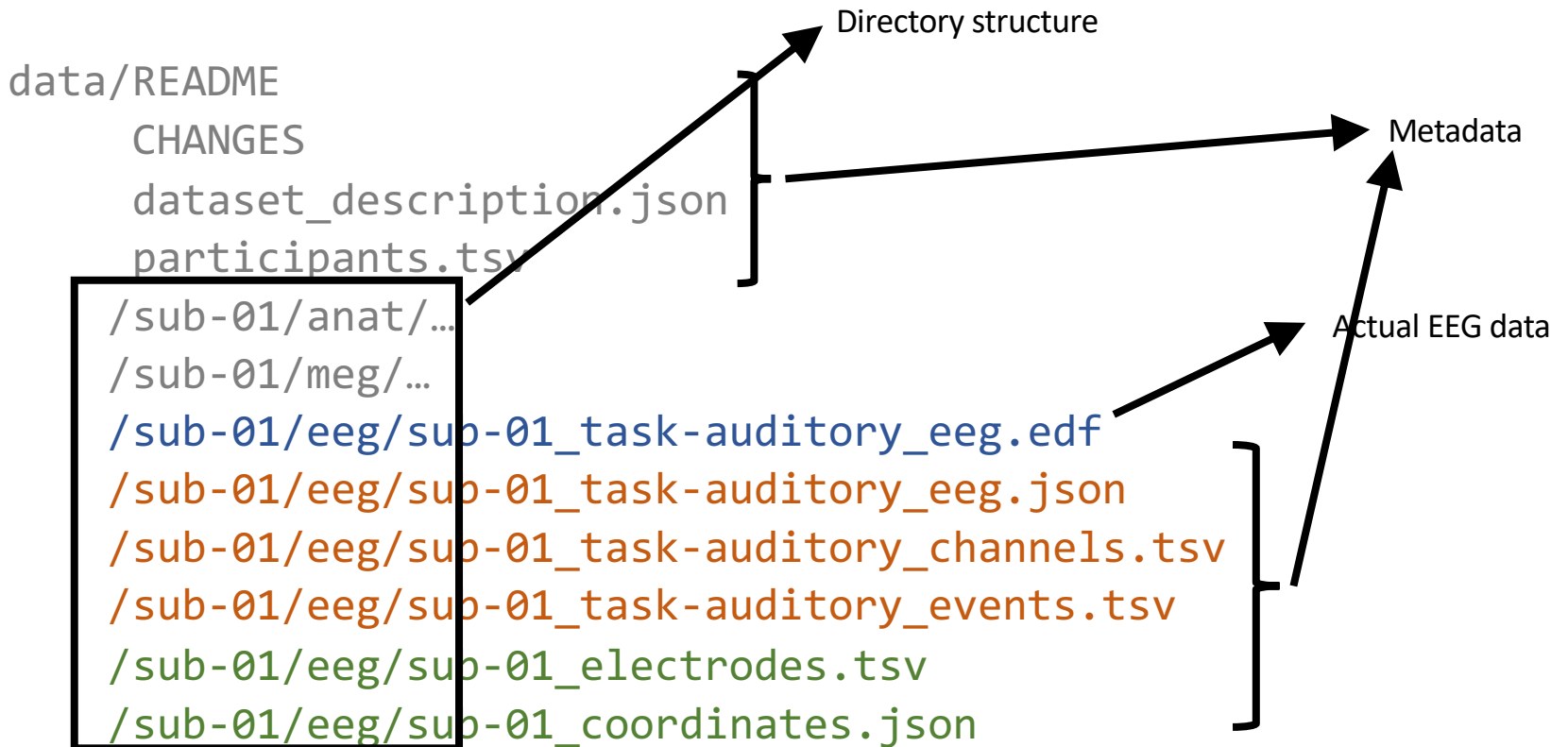
A search engine

A data sharing platform



# BIDS for MRI, MEG, EEG, iEEG ...

in future also for motion capture, eye-tracker, etc.



# BIDS sidecar files for metadata

see also <https://github.com/bids-standard/bids-examples>

- 1) represent otherwise missing data
- 2) make it easier to query/search

As example for EEG:

`_participants.tsv` and json  
`_sessions.tsv` and json  
`_scans.tsv` and json

`_eeg.json`  
`_channels.tsv` and json  
`_electrodes.tsv` and json  
`_coordinates.json`  
`_photos.jpg`

```
{  
  "TaskName": "matchingpennies",  
  "TaskDescription": "The task is emulating a game of 'matching pennies'",  
  "SamplingFrequency": 5000,  
  "Manufacturer": "Brain Products",  
  "ManufacturersModelName": "BrainAmp DC",  
  "CapManufacturer": "Brain Products",  
  "CapManufacturersModelName": "actiCAP 64Ch Standard-2",  
  "EEGChannelCount": 10,  
  "EOGChannelCount": 0,  
  "ECGChannelCount": 0,  
  "EMGChannelCount": 0,  
}
```

name	type	units	status	status_description
FC5	EEG	uV	bad	Contains high frequency noise
FC1	EEG	uV	good	n/a
C3	EEG	uV	good	n/a
CP5	EEG	uV	good	n/a
CP1	EEG	uV	good	n/a
FC2	EEG	uV	good	n/a
FC6	EEG	uV	bad	Low correlation with other channels
C4	EEG	uV	good	n/a
CP2	EEG	uV	good	n/a
CP6	EEG	uV	good	n/a



# BIDS conceptual principles

Data and metadata should be organized in a human- and machine-readable format

Reuse existing data and metadata standards where possible

NIFTI, EDF, NWB

DICOM, CogPo, SI units

Metadata should be human-readable, hence TSV and JSON

Focus on the most common solutions, not the edge cases

One BIDS specification, not separate ones, hence consistent definitions and use of terms across modalities

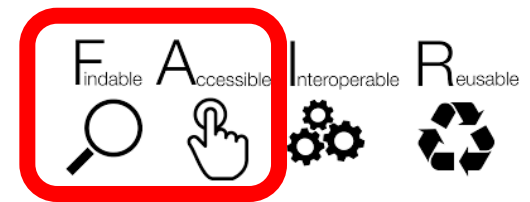
Semantic versioning (version x.y.z), hence backwards incompatible changes must wait until version 2.0

# BIDS technical principles

File naming with some redundance

Inheritance of metadata

“Source data” -> “raw data” -> “derived data”



# BIDS is not a search engine

but it standardizes the metadata

Generic search engines (i.e., web crawlers) will not use BIDS metadata and structure

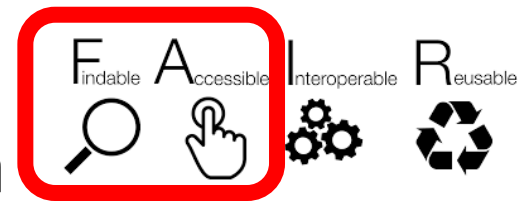
Domain specific search engines might use it

<https://search.datacite.org>

<https://datasetsearch.research.google.com>

<https://www.datalad.org>





# BIDS is not a data sharing platform

So where to share?

## Institutional repository

Donders <https://data.donders.ru.nl>

Radboud University <http://data.ru.nl>

In the UK [Oxford](#), [Cambridge](#), [Edinburg](#)

...

## National repository (in NL)

<https://easy.dans.knaw.nl>

<https://dataverse.nl>

<https://data.4tu.nl>

## Domain specific repository

<https://nemar.org> for EEG etc.

<https://openneuro.org> for all sorts

<https://ebrains.eu> for GDPR

## General repository

<https://zenodo.org>

<https://dataverse.harvard.edu>

<https://osf.io>

## Commercial publishers

<https://datadryad.org>

<https://figshare.com>

# What if you want to use existing BIDS data?

Data sharing platforms often use BIDS

Importing is easy, gives an EEGLAB STUDY structure

Details about the recording and experimental design are included

Other software also uses BIDS

- FieldTrip

- BrainStorm

- MNE-Python

- BrainVision Analyzer

# Example: Wakeman and Henson dataset

***“Multi-subject, multi-modal face processing dataset”***

including MRI, fMRI, MEG and EEG from 18 subjects.

Some history:

- First released as SPM tutorial dataset on an FTP server
- Then published on OpenfMRI as ds000117
- “Data descriptor” paper published in Scientific Data in 2015
- Later BIDSified and published on OpenNeuro as ds000117
- EEG-only version published on OpenNeuro as ds002718

# Example: Wakeman and Henson dataset

Data descriptor paper has already been cited 147 times.

Used in 9 articles with different MEG and/or EEG analysis pipelines in special issue on reproducible group analysis.

EDITORIAL article

Front. Neurosci., 13 April 2022

Sec. Brain Imaging Methods

Volume 16 - 2022 | <https://doi.org/10.3389/fnins.2022.854471>

This article is part of the Research Topic

From raw MEG/EEG to publication: how to perform MEG/EEG group analysis with free academic software

[View all 26 Articles >](#)

## Editorial: From Raw MEG/EEG to Publication: How to Perform MEG/EEG Group Analysis With Free Academic Software



Arnaud Delorme<sup>1,2\*</sup>,



Robert Oostenveld<sup>3,4</sup>,



Francois Tadel<sup>5</sup>,



Alexandre Gramfort<sup>6</sup>,



Srikantan Nagarajan<sup>7</sup> and



Vladimir Litvak<sup>8</sup>

# Example: Wakeman and Henson dataset

<https://openneuro.org/datasets/ds002718/versions/1.0.5>

[https://nemar.org/dataexplorer/detail?dataset\\_id=ds002718](https://nemar.org/dataexplorer/detail?dataset_id=ds002718)



# What if you want to BIDSify your data

Plan for it and start early on, saves you time later

Document experimental details (e.g., trigger codes, acquisition details)

EEGLAB STUDY structure can be exported to BIDS

Other tools for BIDSifying data

- FieldTrip data2bids

- MNE-Python tools

- BIDScoin, Heudiconv (for MRI data)

Upload to <https://nimar.org>, <https://openneuro.org>, or elsewhere

# Doing online computations on BIDS data

NEMAR.org

BrainLife.io

...

