



# ReproNim: A Center for Reproducible Neuroimaging Computation

<http://www.repronim.org>

# Course Material

Screenshot of a GitHub repository page for [Repronim / sfn2018-training](https://github.com/Repronim/sfn2018-training).

The repository contains 103 commits, 3 branches, 1 release, 1 environment, and 10 contributors.

Recent commits:

- jgretke Added FAIR data folder (Latest commit 07b71aa 28 seconds ago)
- .datalad [DATALAD] new dataset (11 days ago)
- FAIR Added FAIR data folder (28 seconds ago)
- Introduction added pdf (4 months ago)
- data Augmented file with MYSTERY (5 months ago)
- section1 Fixed commands for Excercise 3 (5 months ago)
- section23 more github urls tune ups (16 days ago)
- section4
- vm
- .gitattributes [DATALAD] new dataset (11 days ago)
- .gitignore Dump of dumb analysis script for 1st-level FSL GLM (7 months ago)
- README.md Strip away aged/duplicate schedule, point to the website (9 days ago)

<https://github.com/Repronim/sfn2018-training>



# FAIR Data

---

We will provide an introduction to the FAIR Data Principles and strategies for making research outputs that follow these principles - that data is Findable, Accessible, Interoperable and Re-Usable. This section will focus on making data FAIR via BIDS (Brain Imaging Data Structure) and NIDM (Neuroimaging Data Model) data structures, providing an overview of metadata, common data elements, terminologies/ontologies and semantic annotation of data. This section will also introduce pyNIDM, a Python library to manipulate the Neuro Imaging Data Model.

# Section 1

Introduction to FAIR

# What is open data?

---

*“Open data is data that can be freely used, shared and built-on by anyone, anywhere, for any purpose.”*

Open definition;  
Open Knowledge Foundation

# Push for open science-open data

---

“The Administration is committed to ensuring that, to *the greatest extent* and *with the fewest constraints possible* and consistent with law and the objectives set out below, the direct results of federally funded scientific research are made available to and useful for the public, industry, and the scientific community. Such results include peer-reviewed publications and digital data. “

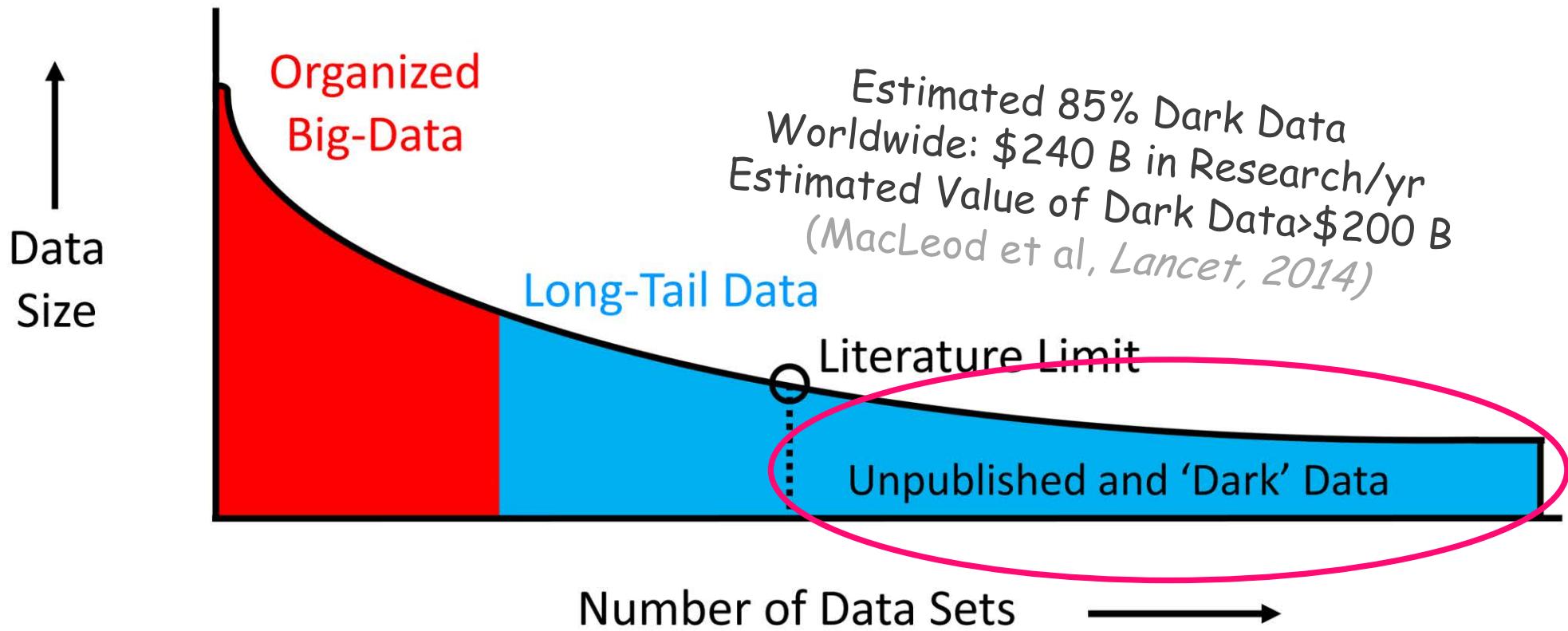
US Office of Science and Technology Policy  
Increasing access to the *results* of federally funded scientific research  
Memo, February 22, 2013

# Why (publish) data?

---

- Increase transparency and reproducibility of research
- Allow and encourage re-use of data sets, e.g., by informaticians, researchers, modelers
- Big data from small data: new insights from analysis across data

# TOWARD 'DATAFICATION' OF BIOMEDICINE

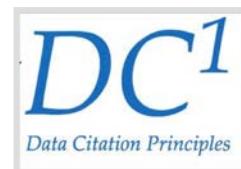


Ferguson et al., 2014, *Nature Neuroscience*

# Data as a Research Product

*“Sound, reproducible scholarship rests upon a foundation of robust, accessible data. For this to be so in practice as well as theory, data must be accorded due importance in the practice of scholarship and in the enduring scholarly record...”*

1. Data should be considered *legitimate, citable products of research*. Data citations should be accorded the same importance in the scholarly record as citations of other research objects, such as publications.
2. Data citations should facilitate giving scholarly credit and normative and legal attribution to all contributors to the data, recognizing that a single style or mechanism of attribution may not be applicable to all data.
3. In scholarly literature, whenever and wherever a claim relies upon data, the corresponding data should be cited.



# An Ecosystem for Data Citation



## On the road to robust data citation

*Scientific Data* 5, Article number: 180095 (2018) | Download Citation ↓

Scientific Data is changing the way we incorporate links into our data citations. We will now be taking advantage of the resolver services offered by identifiers.org and N2T.net to provide more standardized and predictable links for biomedical datasets that have accession identifiers when they are cited in our publications.

These two parallel services help address an important problem in data citation: many of the most important community-supported biomedical repositories use accession identifier systems that must be resolved according to repository-specific linking rules. Indeed, a single repository may offer multiple link structures that all point to the same dataset. An article published today in the journal describes how identifiers.org and N2T.net, provided by the EMBL-EBI and the California Digital Library respectively, are working together to provide machine-resolvable persistent identifiers for datasets across a wide range of biomedical data repositories<sup>1</sup>.



# FAIR Principles

## SCIENTIFIC DATA

OPEN

SUBJECT CATEGORIES  
» Research data  
» Publication characteristics

Comment: The FAIR Guiding Principles for scientific data management and stewardship

Thompson, Johan van der Lei, Enk van Mulligen, Jan Velterop, Andra Waagmeester, Wittneburg, Katherine Wolstencroft, Jun Zhao, and Barend Mons

Open data is about MORE THAN DISCLOSURE it must be Fair

<http://www.nature.com/sdata/>

nature publishing group npg

- Findable
- Accessible
- Interoperable
- Reusable



- **Findable:** A core aspect of an open data ecosystem - facilitate access to a collection of diverse research resources
- **Accessible:** Assist users by directing them to available resources for that researcher
- **Interoperable:** Provide unified views of certain data and information across resources
- **Reusable:** Supports community standards and develops standard representations for resources
- **Attribution and Recognition:** Initiatives to unambiguously cite resources and data

Assist researchers in ensuring that their resources (e.g. software tools, core facilities,...) and data are FAIR

# Making data FAIR

- FAIR data starts with good laboratory management
- Easiest way to make data FAIR is to submit it to a data repository
  - Specialized community repository
  - General repository
  - Institutional repository
- Journals are starting to post lists of repositories that are acceptable (e.g.,  
<http://journals.plos.org/plosone/s/data-availability>)
- Data repositories act in the role of publishers for data:
  - links are stable
  - metadata is consistent
  - community standards are adhered to
- Data journals may have more stringent requirements

The cover of the journal issue features the title 'SCIENTIFIC DATA' in large blue letters, with 'IN PRESS' written diagonally across it. Below the title is the subtitle 'The FAIR Guiding Principles for scientific data management and stewardship'. A list of authors is provided, followed by a detailed description of the FAIR principles. The right side of the cover includes the 'open data' logo and the 'Fair' logo.

**IN PRESS**

**SCIENTIFIC DATA**

**The FAIR Guiding Principles for scientific data management and stewardship**

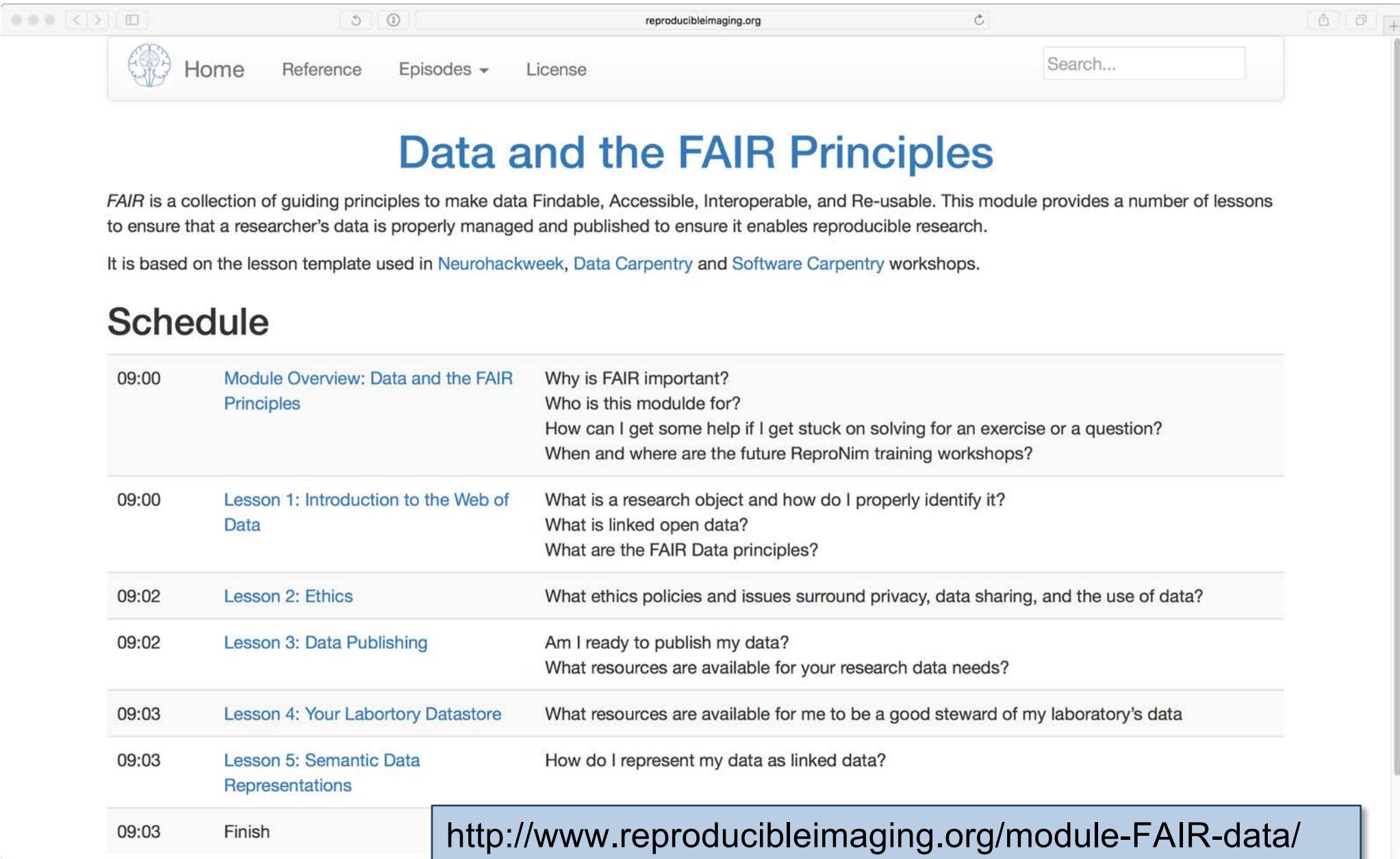
Mark D. Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E Boume, Jildau Bouwman, Anthony J Brookes, Tim Clark, Mercé Crosas, Ingrid Dillo, Olivier Dumon, Scott Edmunds, Chris T Evelo, Richard Finkers, Alejandra Gonzalez-Beltran, Alasdair J G Gray, Paul Groth, Carole Goble, Jeffrey S. Grethe, Jaap Heringa, Peter A.C. 't Hoen, Rob Hooft, Tobias Kuhn, Ruben Kok, Joost Kok, Scott J. Lusher, Maryann E. Martone, Albert Mons, Abel L. Packer, Bengt Persson, Philippe Rocca-Serra, Marco Roos, Rene van Schaik, Susanna-Assunta Sansone, Erik Schultes, Thierry Sengstag, Ted Slater, George Strawn, Morris A. Swertz, Mark Thompson, Johan van der Lei, Erik van Mulligen, Jan Velterop, Andra Waagmeester, Peter Wittenburg, Katherine Wolstencroft, Jun Zhao, and Barend Mons

**open  
data**  
is about  
MORE  
THAN  
DISCLOSURE  
it must be  
**Fair**

- Findable
- Accessible
- Interoperable
- Reusable

<http://www.nature.com/sdata/> nature publishing group npg

# ReproNim Training



The screenshot shows a web browser window with the URL [reproducibleimaging.org](http://reproducibleimaging.org) in the address bar. The page title is "Data and the FAIR Principles". The content area contains a paragraph about FAIR principles, a note about lesson templates, and a schedule table.

**Data and the FAIR Principles**

FAIR is a collection of guiding principles to make data Findable, Accessible, Interoperable, and Re-usable. This module provides a number of lessons to ensure that a researcher's data is properly managed and published to ensure it enables reproducible research.

It is based on the lesson template used in [Neurohackweek](#), [Data Carpentry](#) and [Software Carpentry](#) workshops.

## Schedule

09:00	<a href="#">Module Overview: Data and the FAIR Principles</a>	Why is FAIR important? Who is this modulde for? How can I get some help if I get stuck on solving for an exercise or a question? When and where are the future ReproNim training workshops?
09:00	<a href="#">Lesson 1: Introduction to the Web of Data</a>	What is a research object and how do I properly identify it? What is linked open data? What are the FAIR Data principles?
09:02	<a href="#">Lesson 2: Ethics</a>	What ethics policies and issues surround privacy, data sharing, and the use of data?
09:02	<a href="#">Lesson 3: Data Publishing</a>	Am I ready to publish my data? What resources are available for your research data needs?
09:03	<a href="#">Lesson 4: Your Labtory Datastore</a>	What resources are available for me to be a good steward of my laboratory's data
09:03	<a href="#">Lesson 5: Semantic Data Representations</a>	How do I represent my data as linked data?
09:03	Finish	<a href="http://www.reproducibleimaging.org/module-FAIR-data/">http://www.reproducibleimaging.org/module-FAIR-data/</a>

# ReproNim Training

A screenshot of a web browser window displaying the ReproNim Training website at [reproducibleimaging.org](http://reproducibleimaging.org). The page title is "Data and the FAIR Principles". The navigation bar includes links for Home, Reference, Episodes (with a dropdown menu), License, and a search bar. The main content area features a large section titled "Lesson 1: Introduction to the Web of Data". Below this, there is an "Overview" section with teaching and exercise times, questions, and objectives.



## Data and the FAIR Principles



### Lesson 1: Introduction to the Web of Data

#### Overview

**Teaching:** Self Paced min

**Exercises:** 2 min

#### Questions

- What is a research object and how do I properly identify it?
- What is linked open data?
- What are the FAIR Data principles?

#### Objectives

- Understanding identifiers
- History of open and linked data
- Overview of the FAIR principles

### Introduction

This lesson provides an overview of strategies for making research outputs available through the web, with an emphasis on data. It introduces concepts such as persistent identifiers, linked data, the semantic web and the FAIR principles. It is designed for those with little to no familiarity with these concepts. More technical discussions can be found in the reference materials.

#### Lesson Episodes

- Overview of the current ecosystem
- Research objects and identifier systems
- Short history of open & linked data technologies
- Benefits of linked open data with examples
- Towards the FAIR principles

# Section 2

Interoperable Data and BIDS

# Motivation

- Imaging experiments are complicated and can be organized in many different ways
- Many files/file types per subject
- Members of the same lab may use different ways to arrange data
- Difficulty sharing data within and across large scale projects



"Think this is bad? You should see the inside of my head."

# Benefits of BIDS

- Standardized organization and format makes it easy to share data within/across labs
- More software tools are available that use this structure
- OpenNeuro and others can accept this format
- Validation tools are available to check data integrity
- Focus is on **raw data** (minimally processed), but source (e.g., DICOM) and derived data (e.g., a contrast map, post processed images) can be included as well

# BIDS Resources

## **BIDS Specification**

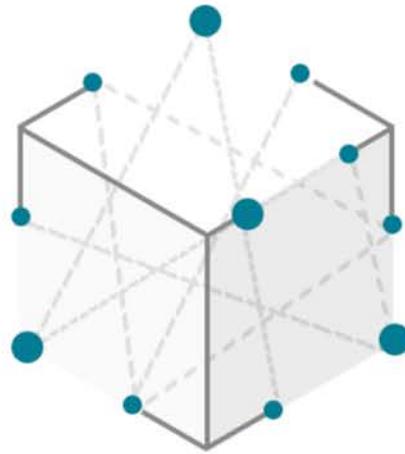
[http://bids.neuroimaging.io/bids\\_spec.pdf](http://bids.neuroimaging.io/bids_spec.pdf)

## **BIDS Starter Kit**

<https://github.com/INCF/bids-starter-kit>

## **BIDS Validator**

<https://github.com/INCF/bids-validator>



# OpenNEURO

A free and open platform for analyzing  
and sharing neuroimaging data

---

<https://openneuro.org>

# Multi-echo fMRI replication sample of autobiographical memory, prospection and theory of mind reasoning tasks

uploaded on 03/26/2018 - 3 months ago

last modified 04/09/2018 - 2 months ago

authored by Elizabeth DuPre, Wen-Ming Luh and R. Nathan Spreng

★ 1   🐾 2   🐞 173   ⬇ 3

**Files: 848, Size: 19.16GB, Subjects: 31, Session: 1**

**Available Tasks** : cuedSGT, rest

**Available Modalities** : bold, T1w

Published   Uploader is Following

## AUTHORS

Elizabeth DuPre

Wen-Ming Luh

<https://openneuro.org/datasets/ds000210/versions/00002>

## BIDS Validation



Valid

2 WARNINGS

## Dataset File Tree

– dataset_description.json
⬇ DOWNLOAD    ⏺ VIEW
– participants.tsv
⬇ DOWNLOAD    ⏺ VIEW
– README
⬇ DOWNLOAD    ⏺ VIEW
– T1w.json
⬇ DOWNLOAD    ⏺ VIEW
– task-cuedSGT_echo-1_bold.json
⬇ DOWNLOAD    ⏺ VIEW
– task-cuedSGT_echo-2_bold.json
⬇ DOWNLOAD    ⏺ VIEW
– task-cuedSGT_echo-3_bold.json
⬇ DOWNLOAD    ⏺ VIEW
– task-rest_echo-1_bold.json
⬇ DOWNLOAD    ⏺ VIEW
– task-rest_echo-2_bold.json
⬇ DOWNLOAD    ⏺ VIEW
– task-rest_echo-3_bold.json
⬇ DOWNLOAD    ⏺ VIEW
– sub-01
– sub-02
– sub-03
– sub-04

# Dataset\_Description.json

The screenshot shows the OpenNEURO interface for a dataset. At the top, there's a navigation bar with links for PUBLIC DASHBOARD, SUPPORT, FAQ, and SIGN IN. Below that, the main content area has a title 'Versions' and two entries: v2 (Apr 9, 2018) and v1 (Mar 26, 2018). A download icon is shown next to the version numbers. The main content area displays the 'dataset\_description.json' file content, which is a JSON object:

```
{  
    "Name": "Multi-echo fMRI replication sample of autobiographical memory, prospection and theory of mind reasoning tasks",  
    "BIDSVersion": "1.0.2",  
    "License": "PDDL",  
    "Authors": [  
        "Elizabeth DuPre",  
        "Wen-Ming Luh",  
        "R. Nathan Spreng"  
    ]  
}
```

# BIDS Specifications

## 7 Directory structure

Overall directories hierarchy is

- sub-<participant\_label>[/ses-<session\_label>]/<data\_type>/
- [code/]
- [derivatives/]
- [stimuli/]
- [sourcedata/]

where [] depicts OPTIONAL content (the same nomenclature is used throughout the spec). Session level is OPTIONAL, first we detail single session example. See below (section 9) for an example with multiple sessions.

### 7.1 Single session example

This is an example of the folder and file structure. Because there is only one session, the session level is not required by the format. For details on individual files see descriptions in the next section:

- **sub-control01**
  - **anat**
    - sub-control01\_T1w.nii.gz
    - sub-control01\_T1w.json
    - sub-control01\_T2w.nii.gz
    - sub-control01\_T2w.json

# BIDS Specifications

## 8.1.1 dataset\_description.json

The file `dataset_description.json` is a JSON file describing the dataset. Every dataset MUST include this file with the following fields:

Name	REQUIRED. Name of the dataset.
BIDSVersion	REQUIRED. The version of the BIDS standard that was used.
License	RECOMMENDED. What license is this dataset distributed under? The use of license name abbreviations is suggested for specifying a license. A list of common licenses with suggested abbreviations can be found in Appendix II.
Authors	OPTIONAL. List of individuals who contributed to the creation/curation of the dataset.
Acknowledgements	OPTIONAL. List of individuals who contributed to the creation/curation of the dataset.
HowToAcknowledge	OPTIONAL. Instructions how researchers using this dataset should acknowledge the original authors. This field can also be used to define a publication that should be cited in publications that use the dataset.

# BIDS Specifications

---

Example:

```
{
```

```
  "Name": "The mother of all experiments",
```

```
  "BIDSVersion": "1.0.1",
```

```
  "License": "CC0",
```

```
  "Authors": ["Paul Broca", "Carl Wernicke"],
```

```
  "Acknowledgements": "Special thanks to Korbinian Brodmann for help in formatting this dataset in BIDS. We thank Alan Lloyd Hodgkin and Andrew Huxley for helpful comments and discussions about the experiment and manuscript; Hermann Ludwig Helmholtz for administrative support; and Claudius Galenus for providing data for the medial-to-lateral index analysis.",
```

```
  "HowToAcknowledge": "Please cite this paper:
```

```
https://www.ncbi.nlm.nih.gov/pubmed/001012092119281",
```

```
  "Funding": ["National Institute of Neuroscience Grant F378236MFH1", "National Institute of Neuroscience Grant 5RMZ0023106"],
```

```
  "ReferencesAndLinks": ["https://www.ncbi.nlm.nih.gov/pubmed/001012092119281",  
"Alzheimer A., & Kraepelin, E. (2015). Neural correlates of presenile dementia in humans. Journal of Neuroscientific Data, 2, 234001. http://doi.org/10.1038/jnd.2015.7"],
```

```
  "DatasetDOI": "10.0.2.3/dfjj.10"
```

```
}
```

---

[http://bids.neuroimaging.io/bids\\_spec.pdf](http://bids.neuroimaging.io/bids_spec.pdf)

# Exercise #1

## Basic BIDS



# Basic BIDS Data (#1)

## Step 1: Setup work folder

```
vagrant@nitricce:~$ mkdir section1  
vagrant@nitricce:~$ cd section1
```

## Step 2: Copy anatomical images

```
vagrant@nitricce:~$ cp ~/workspace/Indiv_Diffs_ReadingSkill/sub-01/anat/* .  
vagrant@nitricce:~$ cp ~/workspace/Indiv_Diffs_ReadingSkill/sub-02/anat/* .  
...  
vagrant@nitricce:~$ mkdir Repro_BIDS
```

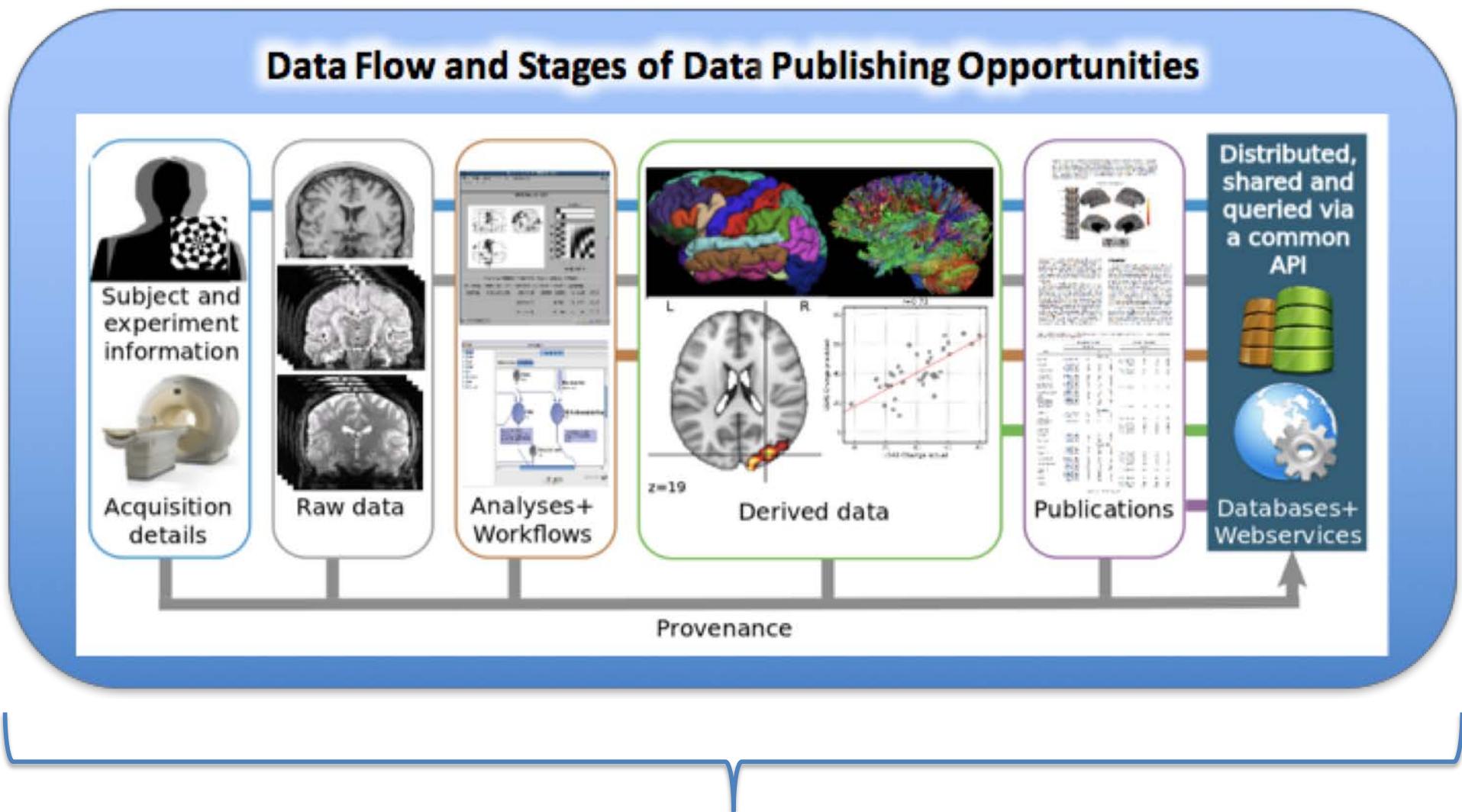
## Step 3: Create a structured BIDS dataset including your dataset description



# Section 3

FAIR Metadata and CDEs

# General Neuroimaging Workflow



**Metadata Descriptions**

# BIDS Specifications

---

## 8.9 Participant file

Template:

(single session case)

`participants.tsv`

`participants.json`

`phenotype/<measurement_tool_name>.tsv`

`phenotype/<measurement_tool_name>.json`

Optional: Yes

The purpose of this file is to describe properties of participants such as age, handedness, sex, etc. In case of single session studies this file has one compulsory column `participant_id` that consists of `sub-<participant_label>`, followed by a list of optional columns describing participants. Each participant needs to be described by one and only one row.

### 8.9.1 `participants.tsv` example:

<code>participant_id</code>	<code>age</code>	<code>sex</code>	<code>group</code>
<code>sub-control01</code>	34	M	<code>control</code>
<code>sub-control02</code>	12	F	<code>control</code>
<code>sub-patient01</code>	33	F	<code>patient</code>

# Exercise #2

## BIDS Participants.tsv



# Basic BIDS Data (#2)

Step 1: Go to your BIDS folder

```
vagrant@nitricce:~$ cd ~/section1/Repro_BIDS
```

Step 2: Create a participants.tsv file

Step 3: Validate your BIDS dataset

<https://github.com/INCF/bids-validator>



# Valid BIDS Dataset

A screenshot of a Linux desktop environment showing a Firefox browser window. The window title is "BIDS Validator". The address bar shows the URL "bids-standard.github.io/bids-validator/". The main content area displays the "BIDS Validator" interface. A large button labeled "Select a BIDS dataset to validate" is present, with a "Browse..." button and a message "No directory selected.". Below this, a note states: "Note: Selecting a dataset only performs validation. Files are never uploaded." A modal dialog titled "Repro\_BIDS" is open, showing a "Summary" section with a list: "• 4 Files, 12.46MB", "• 2 - Subjects", and "• 1 - Session". It also shows "Available Tasks" and "Available Modalities" sections, both currently empty. A message at the bottom of the dialog says "We found 1 Warning in your dataset." followed by a link "view 1 warning in 1 file". At the bottom of the page, there are links for "Download error log for Repro\_BIDS", "Click to view details on BIDS specification", "If you have any questions please post on Neurostars", and "The source code for the validator can be found [here](#)". A "Screenshot" button is located in the bottom right corner of the dialog. The desktop background shows a purple vertical dock on the left side with various application icons.

BIDS Validator × m Firefox Privacy Notice – × +

bids-standard.github.io/bids-validator/

BIDS Validator

Select a BIDS dataset to validate

Browse... No directory selected.

Note: Selecting a dataset only performs validation. Files are never uploaded.

Repro\_BIDS

Summary

- 4 Files, 12.46MB
- 2 - Subjects
- 1 - Session

Available Tasks

Available Modalities

- T1w

We found 1 Warning in your dataset.

view 1 warning in 1 file

[Download error log for Repro\\_BIDS](#)

[Click to view details on BIDS specification](#)

If you have any questions please post on [Neurostars](#)

The source code for the validator can be found [here](#)

Screenshot

# Participants.tsv

OpenNEURO

PUBLIC DASHBOARD SUPPORT FAQ SIGN IN

Versions < >

Multi-echo fMRI replication sample of autobiographical memory, prospection and theory of mind reasoning tasks - participants.tsv

PARTICIPANTS.TSV [DOWNLOAD](#)

participant_id	gender	age	physioSampling	restAcquisition
sub-01	M	26	50	after_cuedSGT
sub-02	M	21	50	after_cuedSGT
sub-03	M	22	50	after_cuedSGT
sub-04	M	23	50	after_cuedSGT
sub-05	M	21	50	before_cuedSGT
sub-06	M	19	50	before_cuedSGT
sub-07	F	18	50	before_cuedSGT
sub-08	F	21	50	before_cuedSGT
sub-09	M	20	50	before_cuedSGT
sub-10	F	21	50	before_cuedSGT
sub-11	F	20	50	before_cuedSGT
sub-12	M	21	50	before_cuedSGT
sub-13	F	31	50	before_cuedSGT
sub-14	M	22	50	before_cuedSGT
sub-15	M	22	50	before_cuedSGT
sub-16	M	25	50	before_cuedSGT

# Terms versus Common Data Elements

Age ✓  
http://uri.interlex.org/base/lix\_0100400

PreferredId: PATO:0000011 Type: term OWL Equivalent: owl:Class  
A time quality inhering in a bearer by virtue of how long it has existed.

Version: 2

General Children Relationships Annotations Referenced By

Superclasses: ✓ Expand list Age

Synonyms: Synonym

Existing IDs:

Preferred CURIE	IRI
NLXWIKI:PATO_0000011	http://neurolex.org/wiki/PATO_0000011
obo/PATO_0000011	
elix_0100400	

interview\_age (Age Differentiation Test) ✓  
http://uri.interlex.org/base/lix\_0345073

PreferredId: ILX:0345073 Type: cde OWL Equivalent: owl:Class  
Age in months at the time of the interview/test/sampling/imaging.

Version: 2

General Children Relationships Annotations Referenced By

Superclasses: ✓ Expand list interview\_age (Age Differentiation Test)  
Age Differentiation Test  
NIMH Data Archive  
Common Data Element

Synonyms: Synonym

Existing IDs:

Preferred CURIE	IRI
ILX:0345073	http://uri.interlex.org/base/lix_0345073

Participant Age ✓  
http://uri.interlex.org/base/lix\_0115067

PreferredId: ILX:0115067 Type: cde OWL Equivalent: owl:Class  
Age at the time of study enrolment, expressed in number of years completed at the last birthday.

Version: 1

General Children Relationships Annotations Referenced By

Superclasses: ✓ Expand list Participant Age  
National Cancer Institute  
Common Data Element  
Common data element  
Standard specification  
Narrative resource

Synonyms: Synonym

Existing IDs:

Preferred CURIE	IRI
ILX:0115067	http://uri.interlex.org/base/lix_0115067

# Common Data Elements

Home / Term Dashboard / ILX:0115067

Term View

Participant Age   
[http://uri.interlex.org/base/ilx\\_0115067](http://uri.interlex.org/base/ilx_0115067)



 Suggest term to community

 Back to search results  
 Add new term

PreferredId: ILX:0115067 Type: cde OWL Equivalent: owl:Class  
Age at the time of study enrollment, expressed in number of years completed at the last birthday.

Export:   

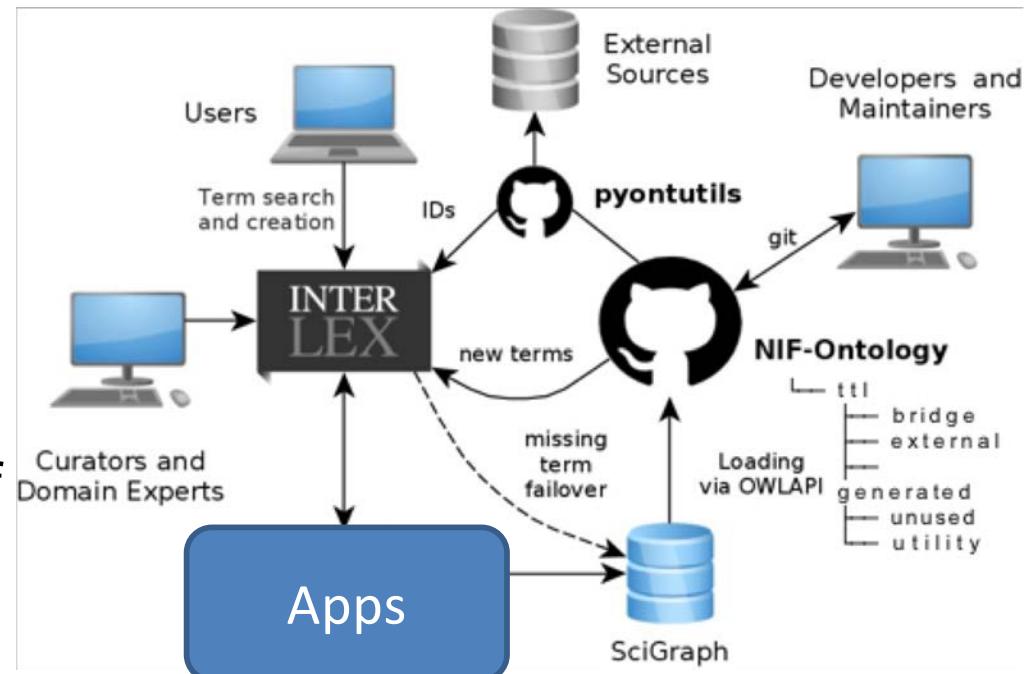
Version: 1

General	Children	Relationships	Annotations	Referenced By														
<p><b>Superclasses:</b></p> <p> Participant Age ↓ National Cancer Institute Common Data Element ↓ Common data element ↓ Standard specification</p>	<p><b>Synonyms:</b></p> <table border="1"><thead><tr><th>Synonym</th><th>Type</th></tr></thead><tbody><tr><td>Age</td><td></td></tr><tr><td>Patient Age</td><td></td></tr><tr><td>Subject Age</td><td></td></tr></tbody></table>	Synonym	Type	Age		Patient Age		Subject Age		<p><b>Existing Ids:</b></p> <table border="1"><thead><tr><th>Preferred</th><th>CURIE</th><th>IRI</th></tr></thead><tbody><tr><td> ILX:0115067</td><td></td><td><a href="http://uri.interlex.org/base/ilx_0115067">http://uri.interlex.org/base/ilx_0115067</a></td></tr></tbody></table>	Preferred	CURIE	IRI	 ILX:0115067		<a href="http://uri.interlex.org/base/ilx_0115067">http://uri.interlex.org/base/ilx_0115067</a>		
Synonym	Type																	
Age																		
Patient Age																		
Subject Age																		
Preferred	CURIE	IRI																
 ILX:0115067		<a href="http://uri.interlex.org/base/ilx_0115067">http://uri.interlex.org/base/ilx_0115067</a>																
<p><b>OWL Equivalent:</b> owl:ObjectProperty</p> <p> Show withdrawn relationships</p> <table border="1"><thead><tr><th>Edit</th><th>Withdrawn</th><th>Term 1</th><th>Relationship type</th><th>Term 2</th><th>Votes</th></tr></thead><tbody><tr><td></td><td> Participant Age</td><td></td><td>Records Instance Of</td><td>Age</td><td>+ 0   - 0</td></tr></tbody></table>					Edit	Withdrawn	Term 1	Relationship type	Term 2	Votes		 Participant Age		Records Instance Of	Age	+ 0   - 0		
Edit	Withdrawn	Term 1	Relationship type	Term 2	Votes													
	 Participant Age		Records Instance Of	Age	+ 0   - 0													

# InterLex

## Community Managed Terminologies

- Dynamic lexicon of biomedical terms
- Constructed to help improve the way scientists communicate about their data
- Provide more powerful means of integrating data across distributed resources.
- InterLex allows for the association of data values (i.e. the value of a field or text within a field) to terminologies enabling the crowdsourcing of data-terminology mappings within and across communities.



Built on the foundation of NeuroLex

<https://interlex.org>

# Term Search



Center for Reproducible Neuroimaging Computation

ABOUT

COMMUNITY RESOURCES

MORE RESOURCES

LITERATURE

MY ACCOUNT 3

Home / Term Dashboard / Term Search

Term Search

echo time



Search Term: echo time Result 1 through 100 of 304.

Search community terms only

First Previous 1 2 3 4 Next Last

## Echo Time

Preferred Id: NLX:149938 Score: 6.374187 Type: term

Time in ms between the middle of the excitation pulse and the peak of the echo produced (kx)

Matches: definition, label

## Effective Echo Time

Preferred Id: NLX:149940 Score: 5.0728955 Type: term

The time in ms between the middle of the excitation pulse and the peak of the echo produced for kx

Matches: definition, label

## TE

Preferred Id: BIRNLEX:2070 Score: 4.9995875 Type: term

Matches: synonyms, annotations

## Duration

Preferred Id: PATO:0001309 Score: 2.1692846 Type: term

A process quality inhering in a bearer by virtue of the bearer's magnitude of the temporal extent be... [\[more\]](#)

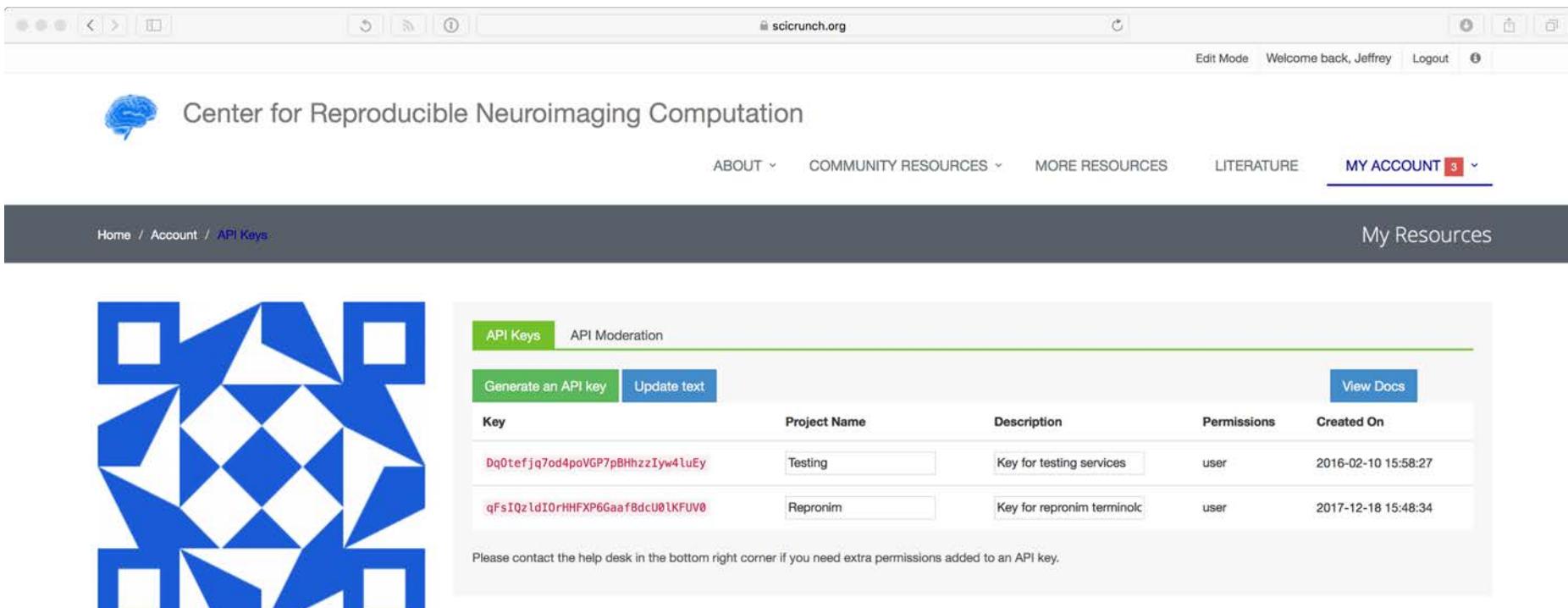
Matches: synonyms, annotations

## Echo Number

Preferred Id: NLX:149935 Score: 1.9904175 Type: term

The echo number used in generating this image. In the case of segmented k-space, it is the effective... [\[more\]](#)

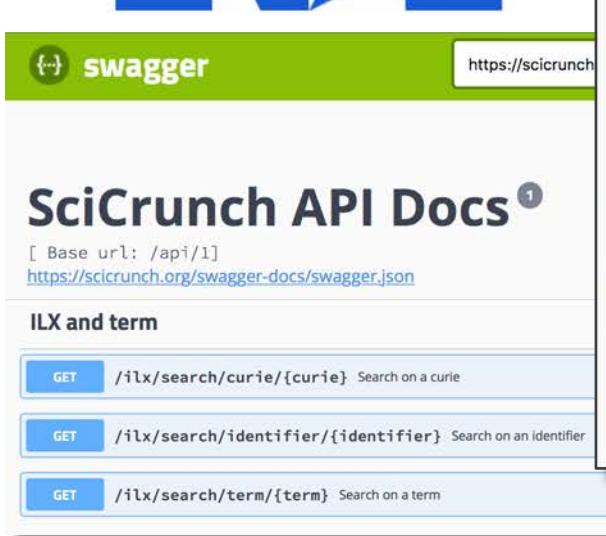
# Access to Services



The screenshot shows the SciCrunch API Keys page. At the top, there's a navigation bar with links for ABOUT, COMMUNITY RESOURCES, MORE RESOURCES, LITERATURE, and MY ACCOUNT (with a red notification badge). Below the navigation is a breadcrumb trail: Home / Account / API Keys. To the right, a "My Resources" section is visible. On the left, there's a large blue geometric logo. The main content area is titled "API Keys" and contains a table with two rows of API key information:

Key	Project Name	Description	Permissions	Created On
Dq0tefjq7od4poVGP7pBHzzIyw4luEy	Testing	Key for testing services	user	2016-02-10 15:58:27
qFsIQzldI0rHHXP6GaafBdcU01KFUV0	Repronim	Key for repronim terminolc	user	2017-12-18 15:48:34

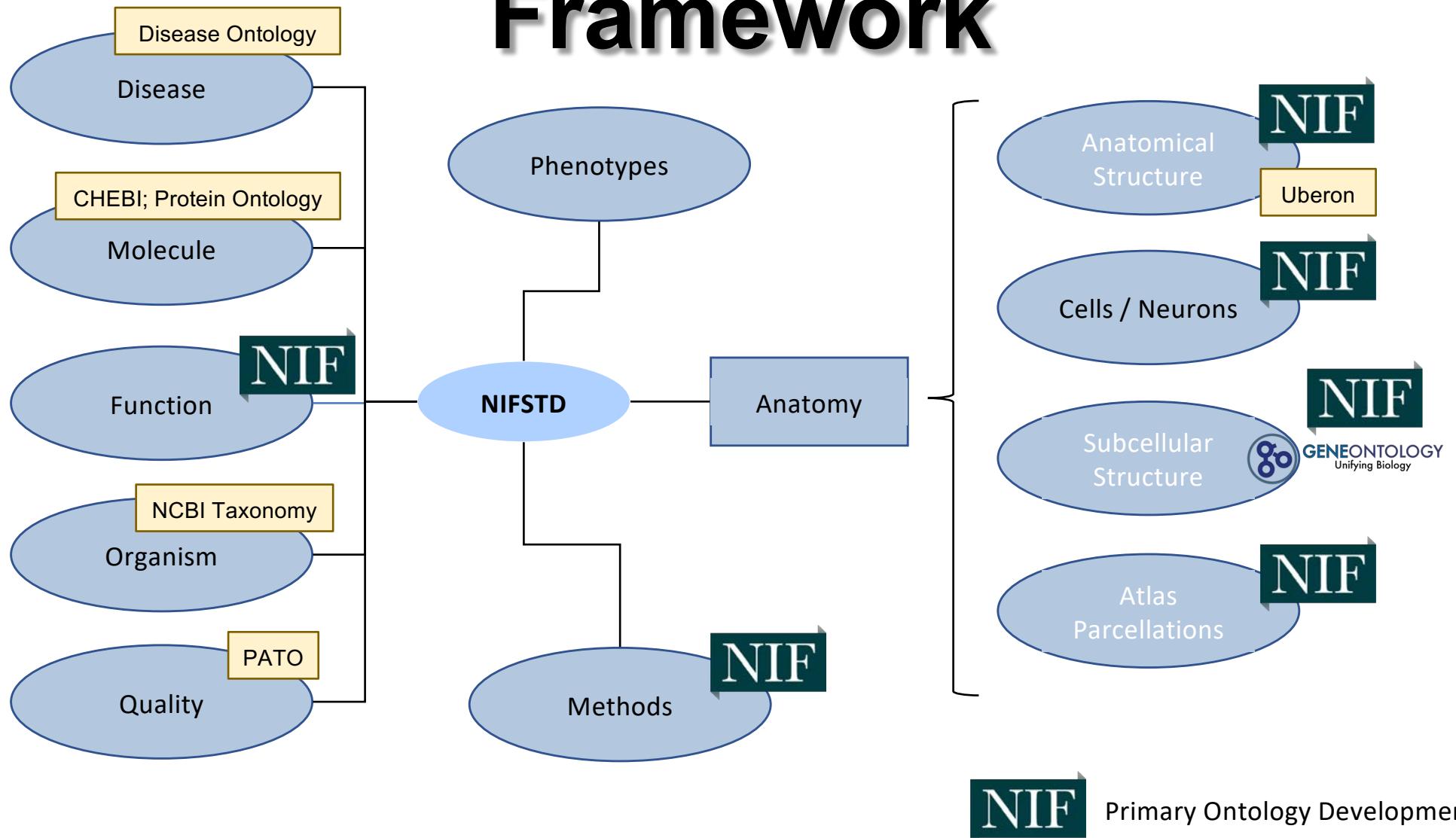
A note at the bottom says: "Please contact the help desk in the bottom right corner if you need extra permissions added to an API key."



The screenshot shows the SciCrunch API Docs page. It features a green header with the "swagger" logo and the URL "https://scicrunch.org". Below the header, the title "SciCrunch API Docs" is displayed with a small "1" indicating updates. A note says "[ Base url: /api/1 ]" and provides the URL "https://scicrunch.org/swagger-docs/swagger.json". The page is divided into sections: "ILX and term", "Elastic", and "Data services". The "Data services" section lists several API endpoints:

- GET /dataservices/literature/pmid retrieve by PMID
- GET /dataservices/federation/search search all data views
- GET /dataservices/literature/search search the Pubmed literature
- GET /dataservices/federation/facets/{viewid} get facets for a data view
- GET /dataservices/federation/data/{viewid} search a single data view

# The NIF Ontologies and Semantic Framework



# Github Management of Ontology

The screenshot shows the GitHub repository page for `SciCrunch / NIF-Ontology`. The repository has 33 issues, 18 stars, and 5 forks. The `Code` tab is selected, displaying the following details:

- 818 commits
- 14 branches
- 4 releases
- 6 contributors
- CC-BY-4.0 license

Recent activity shows a commit by `tgbugs` reserialized using ttiser 1.1.1 on Dec 3, 2017. The commit message is: "docs/external-sources.md added for answers to common questions about the ontology".

The repository contains files such as `docs`, `ttl`, `.gitignore`, `LICENSE`, `README.md`, `catalog-extras`, `dev-reasoner.patch`, `interlex_reserved.txt`, `scicrunch-registry.ttl`, and `README.md`.

**NIF-Ontology**

Repository for the NIF Standard (NIFSTD) Ontologies.

**Notice**

NIFSTD is undergoing significant changes. Please see the [issue tracker](#) for upcoming changes please see the `log` or run `git log` for a quick overview.

**Documentation**

Please see [docs/](#) for documentation about how to develop NIFSTD using this repository.

Please see the [GitHub wiki](#) for documentation about the structure of the ontology and how to load and edit it.

The `scicrunch-registry.ttl` file content is displayed as follows:

```
7726 lines (7042 sloc) | 433 KB
Raw Blame History
@prefix : <file:///ERROR/EMPTY/PREFIX/BANNED/> .
@prefix bearerOf: <http://purl.obolibrary.org/obo/R0_0000053> .
@prefix BFO: <http://purl.obolibrary.org/obo/BFO_> .
@prefix BIRNLLEX: <http://uri.neuinfo.org/nif/nifstd/birnlex_> .
@prefix CHEBI: <http://purl.obolibrary.org/obo/CHEBI_> .
@prefix COGAT: <http://www.cognitiveatlas.org/ontology/cogat.owl#> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .
@prefix ILX: <http://uri.interlex.org/base/ilx_> .
@prefix NIFRID: <http://uri.neuinfo.org/nif/nifstd/readable/> .
@prefix NIFSTD: <http://uri.neuinfo.org/nif/nifstd/> .
@prefix NLX: <http://uri.neuinfo.org/nif/nifstd/nlx_> .
@prefix NLXINV: <http://uri.neuinfo.org/nif/nifstd/nlx_inv_> .
@prefix NLXRES: <http://uri.neuinfo.org/nif/nifstd/nlx_res_> .
@prefix OBI: <http://purl.obolibrary.org/obo/OBI_> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix PATO: <http://purl.obolibrary.org/obo/PATO_> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix xml: <http://www.w3.org/XML/1998/namespace> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

<http://ontology.neuinfo.org/NIF/ttl/NIF-Investigation.ttl> a owl:Ontology ;
    owl:imports <http://ontology.neuinfo.org/NIF/ttl/BIRNLex-OBI-proxy.ttl>,
        <http://ontology.neuinfo.org/NIF/ttl/external/cogat_v0.3.owl>,
        <http://ontology.neuinfo.org/NIF/ttl/external/CogPO.ttl>,
        <http://ontology.neuinfo.org/NIF/ttl/NIF-Government-Granting-Agency.ttl>,
        <http://ontology.neuinfo.org/NIF/ttl/NIF-Resource.ttl>,
        <http://ontology.neuinfo.org/NIF/ttl/NIF-Scientific-Discipline.ttl>,
        <http://ontology.neuinfo.org/NIF/ttl/nif_backend.ttl> ;
```

# BrainVerse

- An **electronic laboratory notebook** built as a cross platform desktop application
- Enables users to **plan experiments, collect, analyze and reuse data, and collaborate**
- **Adds semantic annotation to data** with relevant metadata based on NeuroImaging data Model (NIDM) making experimental neuroimaging study more ***reproducible***, and making data ***FAIR***
- Addresses the limitations of existing Electronic Data Capture systems, such as REDCap and OpenClinica
  - Lack of semantic annotation, do not use underlying graph model, limited query capabilities, do not integrate with other desktop applications, integration and management of specific data types.

The screenshot shows the BrainVerse application interface. At the top, there is a subject profile with fields for 'Subject ID' (S11) and 'Add a Plan' (StudyA). Below this is a table titled 'Session Number' with columns for Session Number, Session Name, Task Name, Instrument Name, and Status. The table contains four rows of data:

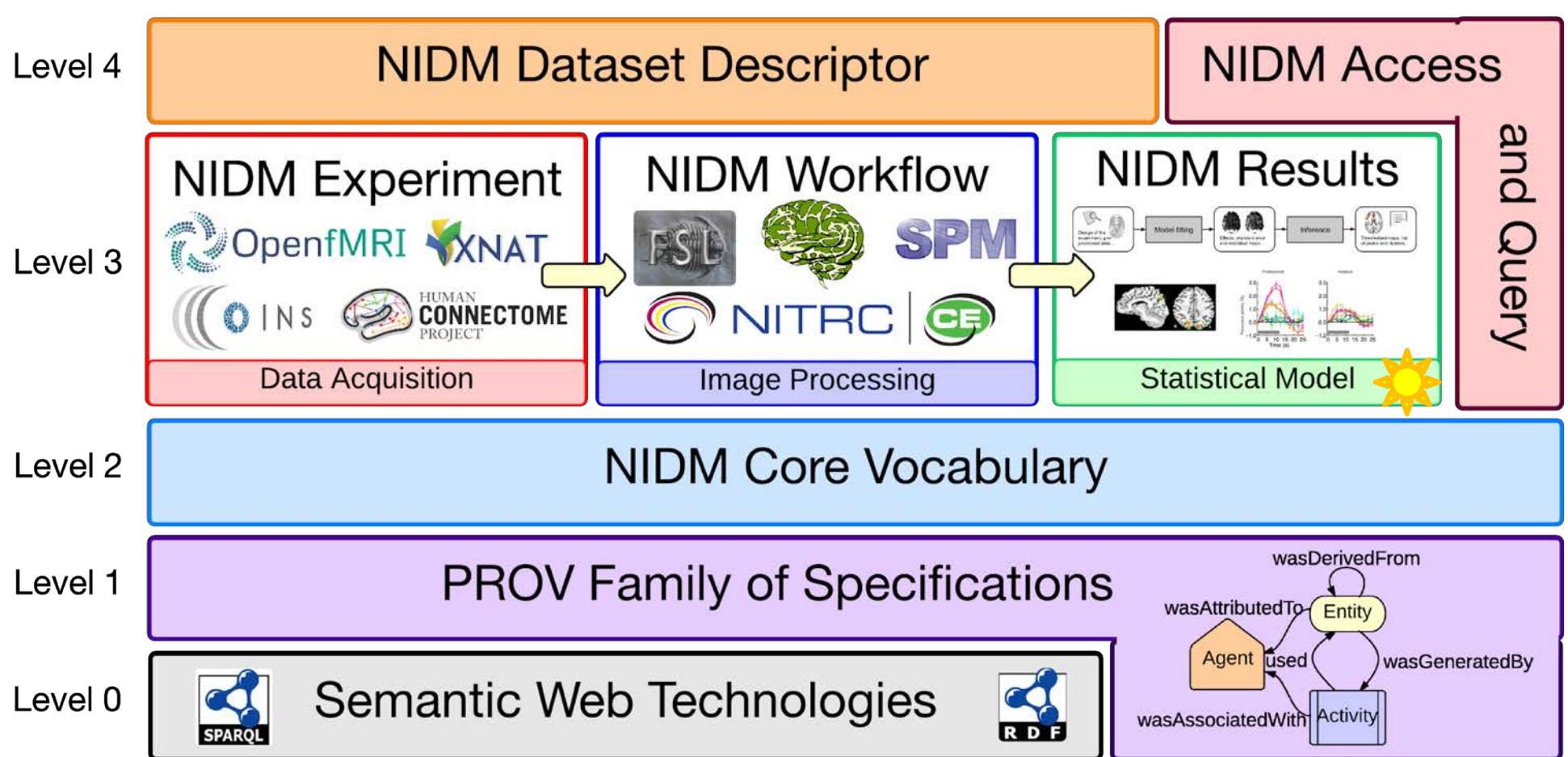
Session Number	Session Name	Task Name	Instrument Name	Status
1	Phone Screen	Call Subjects	terms-bu_demographics01-m32-BUDemographics	Not completed
2	Scan	Scanning	terms-abcd_ra01-m76-scan-list	Not completed
2	Scan	Measure Mindfulness	terms-camm01-m43-camm-edited	Not completed
3	GritScale	Grit Scale Measurement	terms-grit01-m74-12-gritscale	Not completed

<https://github.com/ReproNim/brainverse>

# Section 4

Re-Usable Data and NIDM

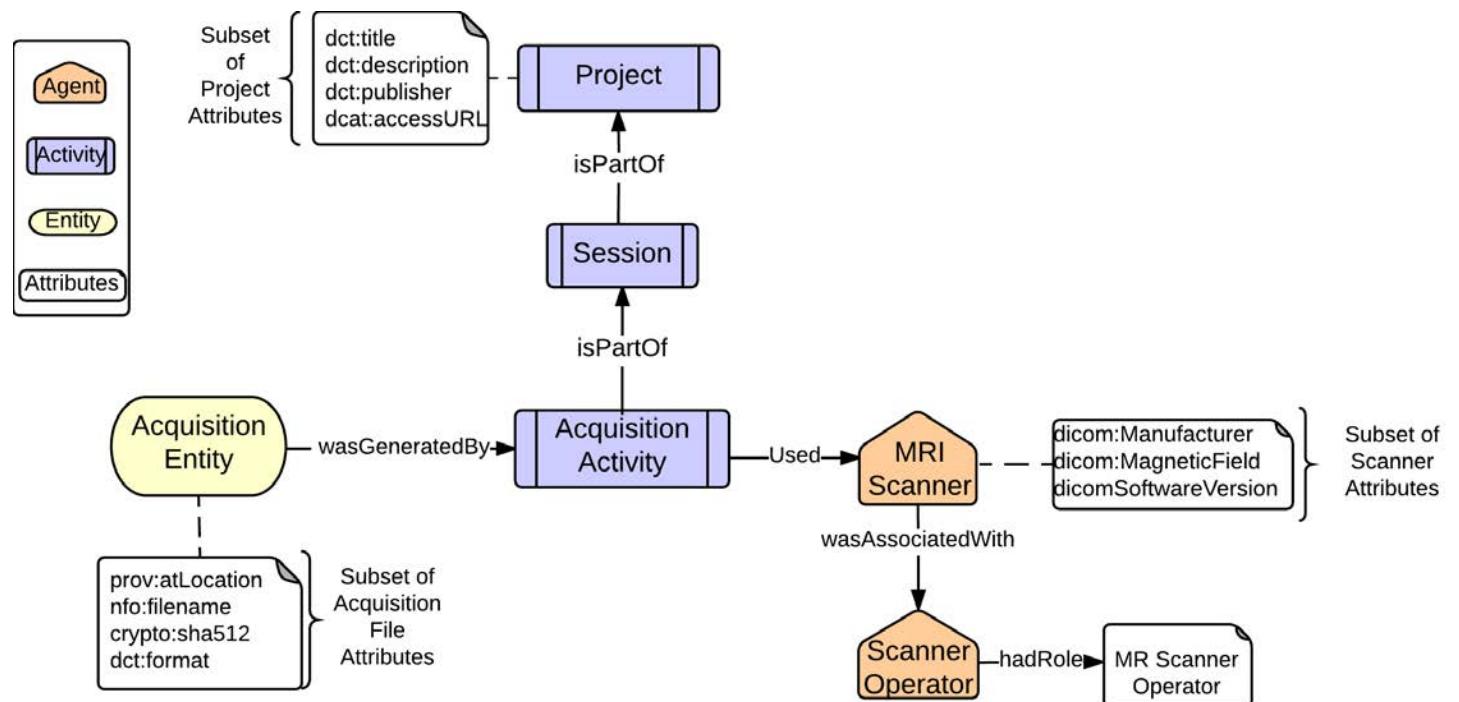
# Neuroimaging Data Model



<http://nidm.nidash.org/>

# NIDM Experiment

- Representation of the experiment design, source data descriptions, and information on the participants and acquisition information.
- Metadata stored using simple Project->Session->Acquisition hierarchy
- Based on W3C Linked Data ([www.w3.org/RDF/](http://www.w3.org/RDF/)) and PROV ([www.w3.org/TR/prov-overview/](http://www.w3.org/TR/prov-overview/)) models.
- Metadata described using subject-predicate-object tuples.
- Tuples composed of namespace:term where “namespace” is web-accessible ontology/terminology containing the precise, unambiguous definition of the “term”.



# PyNIDM

---

[build](#) passing

A Python library to manipulate the [Neuro Imaging Data Model](#).

## Dependencies

---

- [graphviz](#) (native package):
  - Fedora: `dnf install graphviz`
  - OS-X: `brew install graphviz`

## creating a conda environment and installing the library (tested with OSX)

---

- `conda create -n pynidm_py3 python=3 pytest graphviz -y`
- `source activate pynidm_py3`
- `cd PyNIDM`
- `pip install -e .`
- you can try to run a test: `pytest`

## NIDM Experiment Tools

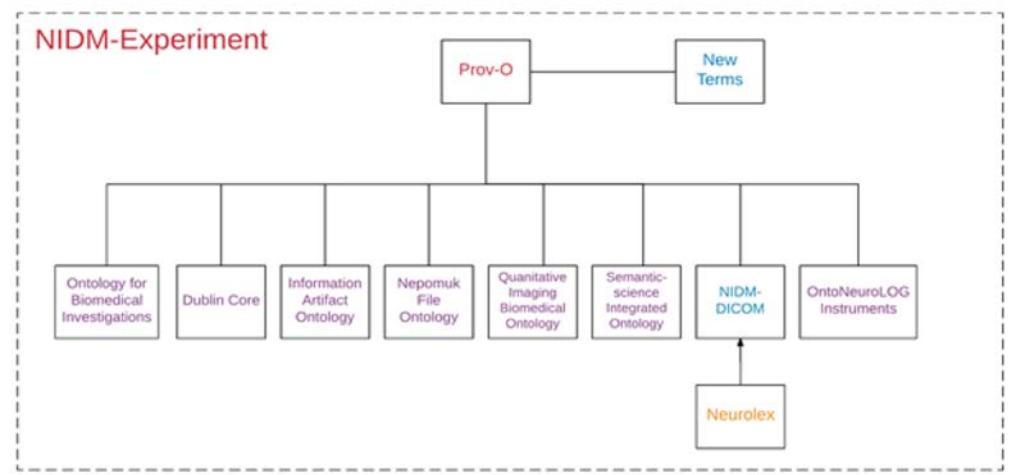
---

### BIDSMRI2NIDM.py

- **Location:** PyNIDM/nidm/experiment/tools/BIDSMRI2NIDM.py
- **Description:** This tool will convert a BIDS MRI directory tree to a NIDM-Experiment document. Currently does not properly handle arbitrary Phenotype files. Will encode in NIDM document but namespace set to BIDS (<http://bids.neuroimaging.io/>) and term to simply variable name from phenotype file which won't de-reference....work in progress....

# PyNIDM – API + Terminology

- Python-based API ([github.com/incf-nidash/PyNIDM](https://github.com/incf-nidash/PyNIDM)) following the simple organizational structure of NIDM-Experiment documents.
  - Functions to create, query, export, import, and transform NIDM-Experiment documents.



- NIDM Terminology / Ontology
  - NIDM ontology reuses existing terms from several other where appropriate definitions exist.
  - NIDM-Experiment ontology: [github.com/incf-nidash/nidm/tree/master/nidm/nidm-experiment/terms](https://github.com/incf-nidash/nidm/tree/master/nidm/nidm-experiment/terms)
  - DICOM-tag ontology: [github.com/incf-nidash/dicom-ontology](https://github.com/incf-nidash/dicom-ontology)

# PyNIDM Terminology Alignment

```
#BIDS Participants file -> NIDM constants      NIDM SUBJECTID =  
mappingsparticipants = {  "participant_id": QualifiedName(provNamespace("ndar",NDAR),  
Constants.NIDM SUBJECTID,  "sex": "src_subject_id")  
Constants.NIDM_GENDER,  "age":  
Constants.NIDM_AGE,  "gender": nidm:_88f65012-e44b-11e7-b9e2-6c4008b8f03e a  
Constants.NIDM_GENDER,  "diagnosis": prov:Agent,  
Constants.NIDM_DIAGNOSIS}  prov:Person ;  
                                         ndar:src_subject_id "sub-10159"^^xsd:string .  
  
  
DICOM =  
Namespace("http://neurolex.org/wiki  
/Category:DICOM_term/")  
  
nidm:_88f8dcec-e44b-11e7-8ba0-  
6c4008b8f03e a sio:file,  
prov:Entity ;  
dicom:EchoTime 3.31e-03 ;  
  
json_keys = {  ##Image terms  "run":  
Constants.NIDM_ACQUISITION_ENTITY,  "ImageType":  
Constants.DICOM["ImageType"],  
"ManufacturerModelName":  
Constants.DICOM["ScanningSequence"],  
"SequenceVariant": Constants.DICOM["SequenceVariant"],  
"ScanOptions": Constants.DICOM["ScanOptions"],  
"MRAcquisitionType":  
Constants.DICOM["MRAcquisitionType"],  
"SequenceName": Constants.DICOM["SequenceName"],  
"RepetitionTime": Constants.DICOM["RepetitionTime"],  
"EchoTime": Constants.DICOM["EchoTime"],
```

# What is Linked Data?

---

“The term Linked Data is used to describe a *method* of exposing, sharing, and connecting data via dereferenceable URIs on the Web.” – Wikipedia

Linked Data is a protocol component of the Semantic Web, using URIs, Real World Objects, Content Negotiation and RDF documents to share information on the web.



# The Semantic Web

- “The semantic web is a vision of information that is understandable by computers, so computers can perform more of the tedious work involved in finding, combining, and acting upon information on the web.” – Wikipedia
- “I have a dream for the Web [in which computers] become capable of analyzing all the data on the Web – the content, links, and transactions between people and computers.”
  - Tim Berners-Lee

# Linked Data Principles

1. Use URIs as names for things
2. Use HTTP URIs so that people can look up (dereference) those names.
3. When someone looks up a URI, provide useful information.
4. Include links to other URIs so that they can discover more things.

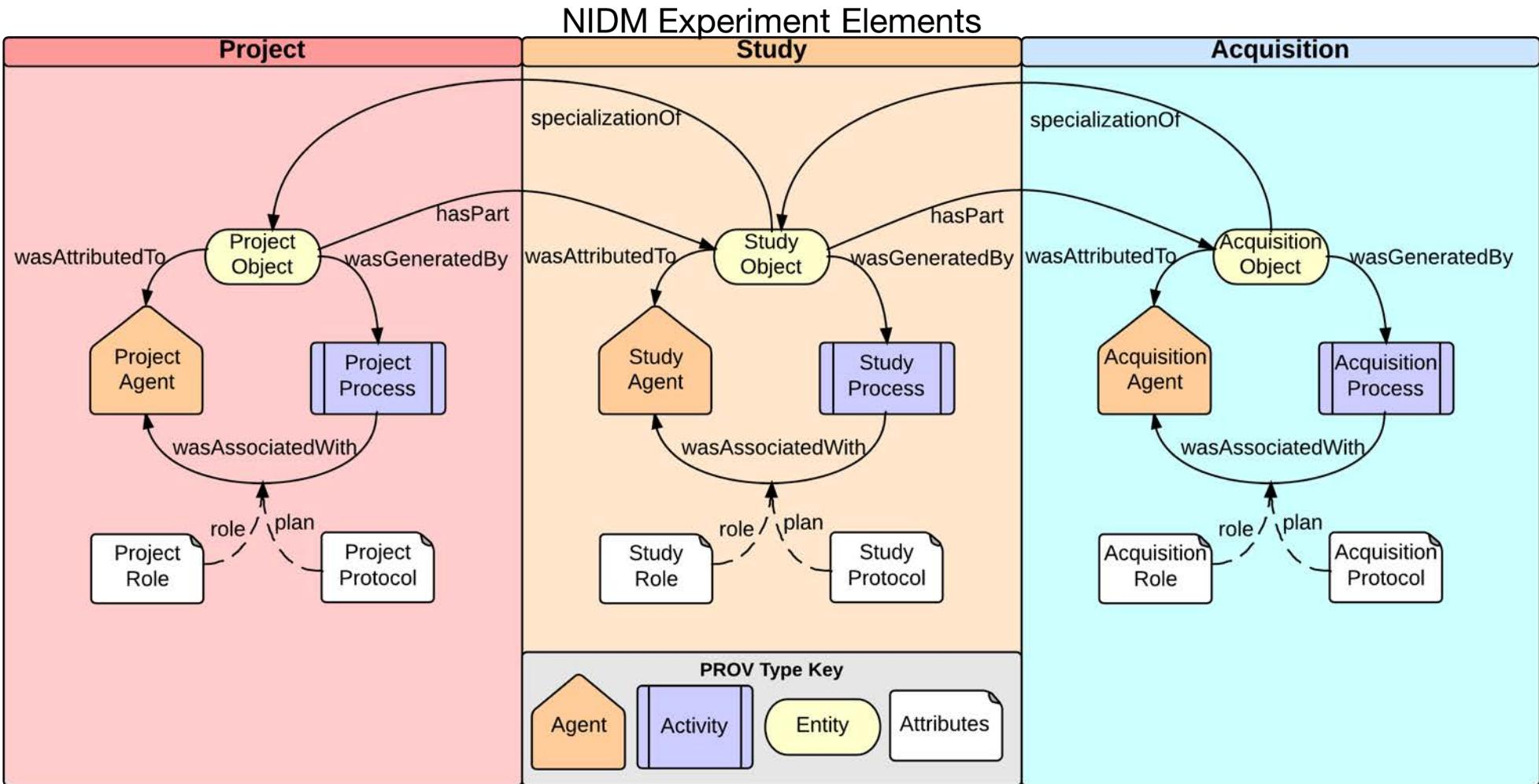
“The Semantic Web isn't just about putting data on the web. It is about making links, so that a person or machine can explore the web of data. With linked data, when you have some of it, you can find other, related, data. Like the web of hypertext, the web of data is constructed with documents on the web.

However, unlike the web of hypertext, where links are relationships anchors in hypertext documents written in HTML, for data they links between arbitrary things described by RDF,. The URIs identify any kind of object or concept. But for HTML or RDF, the same expectations apply to make the web grow.”

# Resource Description Framework (RDF)

- A data model
  - A way to model data
  - i.e. Relational databases use relational data model
- RDF is a triple data model
- Labeled Graph
- Subject, Predicate, Object
- <Cerebellum> <is part of> <Hind Brain>

# NIDM Graph Model



# NIDM Resources

## **NIDM Overview**

<http://nidm.nidash.org/specs/nidm-overview.html>

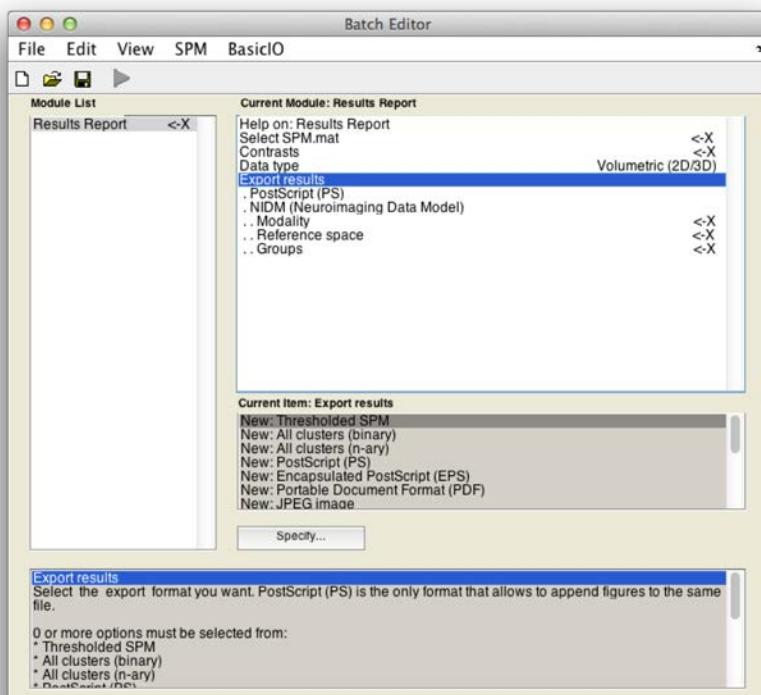
## **NIDM Primer**

<http://nidm.nidash.org/specs/nidm-primer.html>

## **Getting Started with NIDM-Results**

<http://nidm.nidash.org/getting-started/>

# NIDM Results



incf-nidash / nidmresults-fsl

Code Issues 13 Pull requests 1 Projects 1 Wiki Insights

A python library to export FSL's feat results to NIDM-Results <http://nidm.nidash.org/specs/nidm-res...>

SCIENTIFIC DATA

Altmetric: 51 Citations: 3 More detail >

Article | OPEN

## Sharing brain mapping statistical results with the neuroimaging data model

Camille Maumet , Tibor Auer, Alexander Bowring, Gang Chen, Samir Das, Guillaume Flandin, Satrajit Ghosh, Tristan Glatard, Krzysztof J. Gorgolewski, Karl G. Helmer, Mark Jenkinson, David B. Keator, B. Nolan Nichols, Jean-Baptiste Poline, Richard Reynolds, Vanessa Sochat, Jessica Turner & Thomas E. Nichols

Received: 21 June 2016  
Accepted: 21 September 2016  
Published online: 06 December 2016

doi:10.1038/sdata.2016.102  
[Download Citation](#)

Medical research Research data

# **Exercise #3**

## **BIDS to NIDM**



# Setup PyNIDM Environment

## Step 1: Setup Conda environment

```
vagrant@nitricce:~$ cd ~  
vagrant@nitricce:~$ source activate section1
```

## Step 2: Convert BIDS to NIDM

```
vagrant@nitricce:~$ cd ~/section1/Repro_BIDS/  
vagrant@nitricce:~$ ~/PyNIDM/bin/BIDSMRI2NIDM.py -d  
~/section1/Repro_BIDS
```

## Step 3: Review nidm.ttl file



# NIDM.TTL

reprotraining\_2018JUN16 [Running]

Terminal

```
vagrant@nitrccce: ~
  ncicb:Age "0.556"^^xsd:string ;
  prov:wasGeneratedBy nidm:_61ee85c6-717f-11e8-b46b-0800277be888 .

nidm:_61ee85cb-717f-11e8-b46b-0800277be888 a onli:assessment-instrument,
  nidm:AcquisitionObject,
  prov:Entity ;
  bids:ctopp.pa.comp "1.198531065"^^xsd:string ;
  bids:ctopp.rln.raw "-1.104109688"^^xsd:string ;
  bids:mcdecode "1.514966059"^^xsd:string ;
  bids:mcpassfluency "1.680192391"^^xsd:string ;
  bids:mrt "1.358940694"^^xsd:string ;
  bids:piat.l.raw "0.766663412"^^xsd:string ;
  bids:piat.r.raw "1.07060038"^^xsd:string ;
  bids:ppvt.raw "1.303491936"^^xsd:string ;
  bids:sspan2.raw "1.876107955"^^xsd:string ;
  bids:wasi.matr.raw "1.413554228"^^xsd:string ;
  ncicb:Age "1.416"^^xsd:string ;
  prov:wasGeneratedBy nidm:_61ee85ca-717f-11e8-b46b-0800277be888 .

nidm:_61ee85ce-717f-11e8-b46b-0800277be888 a nidm:AcquisitionObject,
  prov:Entity ;
  nidm:AcquisitionModality nidm:MagneticResonanceImaging ;
  nidm:hasImageContrastType nidm:T1Weighted ;
  nidm:hasImageUsageType nidm:Anatomical ;
  nfo:filename "/sub-01/anat/sub-01_T1w.nii.gz"^^xsd:string ;
  prov:wasGeneratedBy nidm:_61ee85cd-717f-11e8-b46b-0800277be888 .

nidm:_61ee85d0-717f-11e8-b46b-0800277be888 a nidm:AcquisitionObject,
  prov:Entity ;
  nidm:AcquisitionModality nidm:MagneticResonanceImaging ;
  nidm:hasImageContrastType nidm:T1Weighted ;
  nidm:hasImageUsageType nidm:Anatomical ;
```



# Exercise #4

## SPARQL Query



# Basic BIDS Data

## Step 1: Locate RDFLIB Qeury File

```
vagrant@nitricce:~$ cd ~/nidm-training  
vagrant@nitricce:~$ ls
```

## Step 2: Review rdf-age-query.py

```
vagrant@nitricce:~$ more rdf-age-query.py
```



# RDFLIB SPARQL Query

```
import rdflib

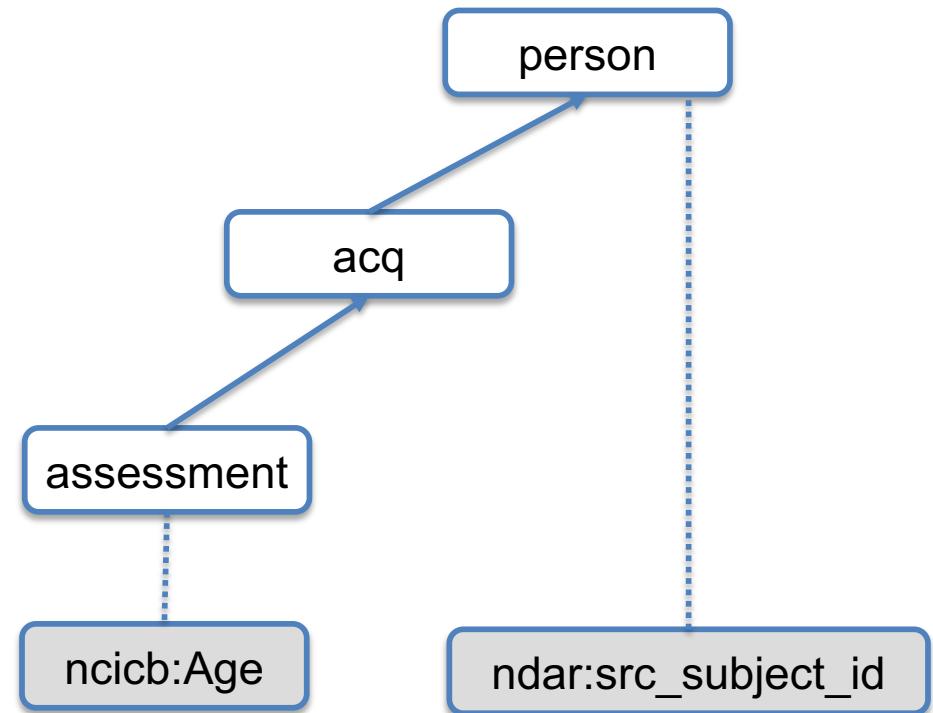
from argparse import ArgumentParser

parser = ArgumentParser()
# parse command line arguments
parser.add_argument('-nidm', dest='nidm_file', required=True, help="NIDM-Exp RDF File to import")
args = parser.parse_args()

g=rdflib.Graph()
g.parse(args.nidm_file, format='ttl')

qres = g.query(
    """SELECT DISTINCT ?id ?age ?assessment
    WHERE {
        ?assessment prov:wasGeneratedBy ?acq .
        ?acq prov:wasAssociatedWith ?person .
        ?assessment ncicb:Age ?age .
        ?person ndar:src_subject_id ?id
    }"""
)

for row in qres:
    print("%s - %s - %s" % row)
```



# Run SPARQL Query

## Step 1: Locate RDFLIB Qeury File

```
vagrant@nitricce:~$ cd ~/nidm-training  
vagrant@nitricce:~$ ls
```

## Step 2: Run Query

```
vagrant@nitricce:~$ python rdf-age-query.py –nidm  
~/section1/Repro_BIDS/nidm.ttl
```

## Experiment:

**FILTER (xsd:float(?age) >= 20)**





# Solution - Exercise #1

## Basic BIDS Data

```
mkdir section1
```

```
cd section1
```

```
cp ~/workspace/Indiv_Diffs_ReadingSkill/sub-01/anat/* .
```

```
cp ~/workspace/Indiv_Diffs_ReadingSkill/sub-02/anat/* .
```

```
ls
```

```
mkdir Repro_BIDS
```

```
cd Repro_BIDS
```

```
mkdir sub-01
```

```
mkdir sub-01/anat
```

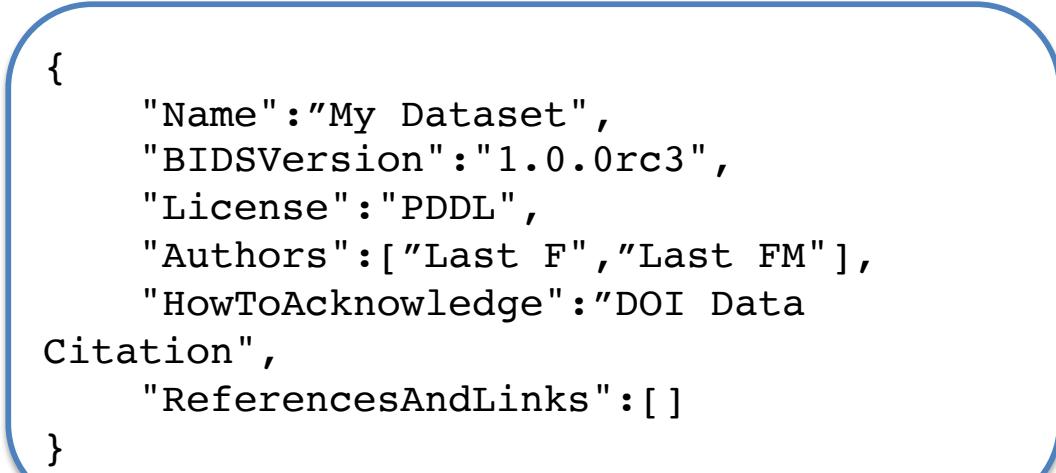
```
cp ../sub-01_T1w.nii.gz ./sub-01/anat
```

```
mkdir sub-02
```

```
mkdir sub-02/anat
```

```
cp ../sub-02_T1w.nii.gz ./sub-02/anat
```

```
emacs dataset_description.json
```

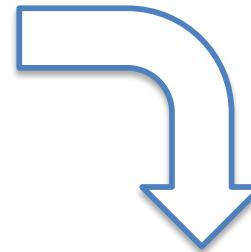


```
{  
    "Name": "My Dataset",  
    "BIDSVersion": "1.0.0rc3",  
    "License": "PDDL",  
    "Authors": ["Last F", "Last FM"],  
    "HowToAcknowledge": "DOI Data  
Citation",  
    "ReferencesAndLinks": []  
}
```

# Solution - Exercise #2

## Basic BIDS Data

```
cd ~/section1/Repro_BIDS/  
emacs participants.tsv
```



```
participant_id age  
sub-01 32  
sub-02 24
```

# Solution - Exercise #3

## Setup PyNIDM Environment

```
cd ~  
source activate section1  
cd ~/section1/Repro_BIDS/  
~/PyNIDM/bin/BIDSMRI2NIDM.py -d  
~/section1/Repro_BIDS  
ls  
more nidm.ttl
```

# Solution – Exercise #4

## SPARQL Query

```
cd ~/nidm-training/  
ls  
more rdf-age-query.py  
python rdf-age-query.py -nidm  
~/section1/Repro_BIDS/nidm.ttl
```

```
(section1) vagrant@nitrcce:~/nidm-training$ python rdf-age-query.py -nidm  
~/section1/Repro_BIDS/nidm.ttl  
sub-02 - 24 - http://purl.org/nidash/nidm#_d6c8b9f1-debb-11e8-b0cd-0800277be888  
sub-01 - 32 - http://purl.org/nidash/nidm#_d6c8b9ed-debb-11e8-b0cd-0800277be888
```