SPM and Data Sharing

Guillaume Flandin
Wellcome Trust Centre for Neuroimaging
University College London
Wavelab and Reproducible Research

Jonathan B. Buckheit and David L. Donoho

Stanford University, Stanford CA 94305, USA

Abstract

Wavelab is a library of MatLab routines for wavelet analysis, wavelet-packet analysis, cosine-packet analysis and matching pursuit. The library is available free of charge over the Internet. Versions are provided for Macintosh, UNIX and Windows machines.

Wavelab makes available, in one package, all the code to reproduce all the figures in our published wavelet articles. The interested reader can inspect the source code to see exactly what algorithms were used, how parameters were set in producing our figures, and can then modify the source to produce variations on our results. Wavelab has been developed, in part, because of exhortations by Jon Claerbout of Stanford that computational scientists should engage in “really reproducible” research.

“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.”
Reproducible Research

- Internet: *distribution, collaboration*.
- Freeware: *GNU licences, copyleft*.
- Quantitative Programming Environments: *high-level, fourth generation programming languages (MATLAB, R, Python, Julia, …)*

In the future, we can envision that publication in computational sciences will change so that reproducibility is integrated into process. One way to do this would be if journals were fully electronic, and if we adopted hypermedia techniques. Then every computationally-generated figure and every computationally-generated table in an article would become linked to the code and the computational environment that produced the figure. If one were interested in a figure, one would click on it with a mouse, and a new window would instantly appear, containing the code that the author of the article used to create the figure. To reproduce the figure, but perhaps change slightly the settings of the display software (for example, to view a surface from a different 3-d perspective), one would simply edit the code in the window and re-run the code; the figure would be re-computed and re-displayed.
SPM & Reproducible Research

Karl Friston’s DEM toolbox
Computational Nosology and Precision Psychiatry

A Proof of Concept

Karl J. Friston
Neuroimaging publications

- Raw data → Pre-processing → Pre-processed data → Analysis → Results → Publication → Publication

Presentation:

- Reporting guidelines: COBIDAS
- Incomplete statistical results
- Ambiguous/incomplete methods
- Metadata is not machine readable

Slide: C. Maumet
International Collaborative Effort

- INCF Neuroimaging data sharing Task Force (NIDASH)
  - Representing 13 labs
  - Weekly teleconferences, focused workshops, GitHub
  - Open

- Stanford Center for Reproducible Neuroscience

- ReproNim

http://nidm.nidash.org
http://reproducibility.stanford.edu
http://www.reproducibleimaging.org
Brain Imaging Data Structure

NeuroImaging Data Model
Brain Imaging Data Structure (BIDS)

“A simple and intuitive way to organise and describe your neuroimaging and behavioural data.”

Benefits of a common standard:

- Minimised curation
  - Within a lab over time
  - Between labs (collaboration and multi-centre studies)
  - Between public databases (e.g. OpenfMRI)
- Error reduction (automated validation)
- Optimised usage of data analysis software (completely automated analysis workflows)
Brain Imaging Data Structure (BIDS)

http://bids.neuroimaging.io/

K.J. Gorgolewski et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Scientific Data (2016)
Brain Imaging Data Structure (BIDS)

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

---

participant_id  age  sex
Sub-001  34  M
Sub-002  22  F
Sub-003  33  F
Brain Imaging Data Structure (BIDS)
Brain Imaging Data Structure (BIDS)

```json
{
  "RepetitionTime": 2,
  "EchoTime": 0.03,
  "FlipAngle": 78,
  "SliceTiming": [0,1.0325,0.06,...],
  "PhaseEncodingDirection": "j-
}
```
BIDS Validator

http://incf.github.io/bids-validator/

Summary
- 40 Files, 18.42kB
- 13 - Subjects
- 1 - Session

Available Tasks
- rhyme judgment

Available Modalities
- bold
- T1w

Your dataset is not a valid BIDS dataset.

view 1 error in 23 files
view 1 warning in 4 files
BIDS Extensions

- Work in progress:
  - PET / SPECT
  - EEG / MEG
  - Model and hypothesis specifications
  - ...

STM
BIDS Apps

BIDS

BIDS App

Derived data

docker run ... bids/spm /bids /output participant --participant_label 01
docker run ... bids/spm /bids /output group

http://bids-apps.neuroimaging.io/
https://www.docker.com/
http://singularity.lbl.gov/

K.J. Gorgolewski et al. BIDS Apps: improving ease of use, accessibility, and reproducibility of neuroimaging data analysis methods. PLOS Computational Biology (2017)
# Available BIDS Apps

<table>
<thead>
<tr>
<th>App</th>
<th>Version</th>
<th>Issues</th>
<th>Build Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>bids/example</td>
<td>v0.0.4</td>
<td>0</td>
<td>passing</td>
</tr>
<tr>
<td>bids/freesurfer</td>
<td>v0.3.2</td>
<td>0</td>
<td>passing</td>
</tr>
<tr>
<td>bids/ndmg</td>
<td>v0.0.46</td>
<td>0</td>
<td>passing</td>
</tr>
<tr>
<td>bids/BROCCOLI</td>
<td>v1.0.0</td>
<td>0</td>
<td>passing</td>
</tr>
<tr>
<td>bids/FibreDensityAndCrosssection</td>
<td>v0.0.3</td>
<td>0</td>
<td>passing</td>
</tr>
<tr>
<td>bids/SPM</td>
<td>v0.0.5</td>
<td>0</td>
<td>passing</td>
</tr>
<tr>
<td>bids/MRIQC</td>
<td>v0.0.2</td>
<td>0</td>
<td>passing</td>
</tr>
<tr>
<td>bids/QAP</td>
<td>image not found</td>
<td>0</td>
<td>passing</td>
</tr>
<tr>
<td>bids/CPAC</td>
<td>v1.0.1e_13</td>
<td>0</td>
<td>passing</td>
</tr>
<tr>
<td>bids/hyperalignment</td>
<td>image not found</td>
<td>0</td>
<td>passing</td>
</tr>
<tr>
<td>bids/mindboggle</td>
<td>v0.0.1</td>
<td>2</td>
<td>failed</td>
</tr>
<tr>
<td>bids/MPrx3_connectome</td>
<td>latest</td>
<td>0</td>
<td>passing</td>
</tr>
<tr>
<td>bids/rs_signal_extract</td>
<td>v0.1</td>
<td>0</td>
<td>passing</td>
</tr>
<tr>
<td>bids/aa</td>
<td>image not found</td>
<td>0</td>
<td>passing</td>
</tr>
<tr>
<td>bids/niak</td>
<td>latest</td>
<td>3</td>
<td>passing</td>
</tr>
<tr>
<td>bids/opnni</td>
<td>latest</td>
<td>0</td>
<td>passing</td>
</tr>
<tr>
<td>bids/mrprepp</td>
<td>v0.2.0</td>
<td>0</td>
<td>passing</td>
</tr>
<tr>
<td>bids/brainiac-srm</td>
<td>latest</td>
<td>0</td>
<td>passing</td>
</tr>
<tr>
<td>bids/nipnetselines</td>
<td>image not found</td>
<td>0</td>
<td>passing</td>
</tr>
<tr>
<td>bids/HCPpipelines</td>
<td>v2.1.7.0_33</td>
<td>0</td>
<td>passing</td>
</tr>
<tr>
<td>bids/MAGEtbrain</td>
<td>image not found</td>
<td>0</td>
<td>passing</td>
</tr>
<tr>
<td>bids/tracula</td>
<td>v0.6.0.2</td>
<td>0</td>
<td>passing</td>
</tr>
</tbody>
</table>

---

**Center for Reproducible Neuroscience**

Reliable easy to use database with versioning + Robust cutting edge "glass box" methods

Providing access to state of the art methods while incentivizing data sharing

http://bids-apps.neuroimaging.io/
Brain Imaging Data Structure

NeuroImaging Data Model
NIDM: a set of specifications to describe neuroimaging data

- **NIDM Experiment**
  - OpenfMRI
  - XNAT
  - INS
  - HUMAN CONNECTOME PROJECT

- **NIDM Workflow**
  - FSL
  - NITRC
  - SPM

- **NIDM Results**
  - Statistical Model

**NIDM Core Vocabulary**

**PROV Family of Specifications**

**Semantic Web Technologies**

- SPARQL
- RDF
NIDM: a set of specifications to describe neuroimaging data
Statistics: \textit{p-values adjusted for search volume}

<table>
<thead>
<tr>
<th></th>
<th>set-level</th>
<th>cluster-level</th>
<th>peak-level</th>
<th>mm mm mm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(p)</td>
<td>(c)</td>
<td>(p_{FWE\text{-corr}})</td>
<td>(q_{FDR\text{-corr}})</td>
</tr>
<tr>
<td>0.000</td>
<td>12</td>
<td>0.000</td>
<td>0.000</td>
<td>425</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>568</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>31</td>
</tr>
<tr>
<td>0.003</td>
<td>0.001</td>
<td>12</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>0.000</td>
<td>0.096</td>
<td>10</td>
<td>0.001</td>
<td>0.000</td>
</tr>
</tbody>
</table>

| Heights above | \(T = 5.25\), \(p = 0.000\) (0.050) |
| Extent threshold | \(k = 0\) voxels |
| Expected voxels per cluster | \(<k> = 0.753\) |
| Expected number of clusters, \(<c> = 0.07\) |
| FWEp: 5.253, FDRp: 6.320, FWEc: 1, FDRc: 10 |

Degrees of freedom = [1.0, 73.0]
FWHM = 9.9 9.9 8.6 mm mm mm; 3.3 3.3 2.9 (voxels)
Volume: 1901367 = 70421 voxels = 1995.7 reseles
Voxel size: 3.0 3.0 3.0 mm mm mm; (reisel = 31.52 voxels)

The table shows 3 local maxima more than 8.0 mm apart.
NIDM-Results pack: Compressed file containing a NIDM-Results serialisation and some or all of the referenced image data files.

SPM export to NIDM-Results

Statistics: \( p\)-values adjusted for search volume

<table>
<thead>
<tr>
<th>( \rho )</th>
<th>( c )</th>
<th>( p_{\text{FWE}} )</th>
<th>( q_{\text{FDR}} )</th>
<th>( k )</th>
<th>( p_{\text{uncorr}} )</th>
<th>( \rho_{\text{FWE}} )</th>
<th>( q_{\text{FDR}} )</th>
<th>( \tau )</th>
<th>( (z) )</th>
<th>( p_{\text{uncorr}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000</td>
<td>12</td>
<td>0.000</td>
<td>0.000</td>
<td>478</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>16.49</td>
<td>Inf</td>
<td>0.000</td>
</tr>
<tr>
<td>0.000</td>
<td>606</td>
<td>0.000</td>
<td>0.000</td>
<td>16.39</td>
<td>Inf</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>10.78</td>
<td>Inf</td>
</tr>
<tr>
<td>0.000</td>
<td>0.000</td>
<td>33</td>
<td>0.000</td>
<td>7.22</td>
<td>6.25</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
<td>7.09</td>
<td>6.16</td>
</tr>
<tr>
<td>0.000</td>
<td>0.001</td>
<td>13</td>
<td>0.000</td>
<td>7.09</td>
<td>6.16</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
<td>7.09</td>
<td>6.16</td>
</tr>
<tr>
<td>0.000</td>
<td>0.051</td>
<td>4</td>
<td>0.025</td>
<td>6.24</td>
<td>5.57</td>
<td>0.000</td>
<td>0.000</td>
<td>0.007</td>
<td>6.42</td>
<td>5.72</td>
</tr>
<tr>
<td>0.000</td>
<td>0.003</td>
<td>10</td>
<td>0.001</td>
<td>6.24</td>
<td>5.57</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
<td>7.09</td>
<td>6.16</td>
</tr>
<tr>
<td>0.006</td>
<td>0.149</td>
<td>2</td>
<td>0.099</td>
<td>5.45</td>
<td>4.97</td>
<td>0.000</td>
<td>0.000</td>
<td>0.002</td>
<td>5.80</td>
<td>5.25</td>
</tr>
<tr>
<td>0.006</td>
<td>0.149</td>
<td>2</td>
<td>0.099</td>
<td>5.45</td>
<td>4.97</td>
<td>0.000</td>
<td>0.000</td>
<td>0.002</td>
<td>5.80</td>
<td>5.25</td>
</tr>
<tr>
<td>0.015</td>
<td>0.233</td>
<td>1</td>
<td>0.233</td>
<td>5.09</td>
<td>4.64</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
<td>5.43</td>
<td>4.96</td>
</tr>
</tbody>
</table>

Print text table
Extract table data structure
Export to CSV file

Export to NIDM-Results

.Locally
Upload to NeuroVault

.nidm.zip
NIDM-Results Viewer

Thomas Maullin-Sapey & Camille Maumet
https://github.com/incf-nidash/nidmresults-spmhtml
Upload NIDM-Results to NeuroVault

Upload NIDM-Results to NeuroVault

Data Sharing Repositories
In addition to using the 3D neuroimaging tool, researchers producing statistical mapping from MRI and PET studies should consider data sharing. NeuroVault is a public repository of brain activation maps supported by the INCF. You can freely share your data and uploading images by visiting neurovault.org.
From NeuroVault to Neurosynth

NIDM: a set of specifications to describe neuroimaging data
Conclusion

- BIDS
- BIDS Apps

- NIDM-Experiment
- NIDM-Workflow

- NIDM-Results
  - Export
  - Viewer
  - Upload to NeuroVault

openfMRI

spm_BIDS.m
hub.docker.com/r/bids/spm/

NIDM Experiment

NIDM Workflow

NIDM-Results

NeuroVault

spm_provenance.m
spm_results_nidm.m
github.com/incf-nidash/nidmresults-spmhtml

NeuroVault
Acknowledgements


NIDM working group: Tibor Auer, Samir Das, Fariba Fana, Guillaume Flandin, Satra Ghosh, Tristan Glatard, Chris Gorgolewski, Karl Helmer, David Keator, Camille Maumet, Nolan Nichols, Jean-Baptiste Poline, Vanessa Sochat, Jason Steffener, Jessica Turner.

Stanford Center for Reproducible Neuroscience: Oscar Esteban, Chris Gorgolewski, Russ Poldrack.

Warwick neurodata sharing: Tom Nichols, Alex Bowring, Tom Maullin-Sapey, Ruth Pauli, Peter Williams.