The NiPreps ecosystem for reproducible neuroimaging

Russ Poldrack Stanford University

The standard design (anti-)pattern for fMRI preprocessing

- Pick a single software package
 - Usually based on considerations other than performance









The standard design (anti-)pattern for fMRI preprocessing

- Pick a single software package
 - Usually based on considerations other than performance
- String together the tools from that package into a script to run the preprocessing workflow

#!/bin/bash

1
15

subj="sub-\$id" echo cd \$subj

f	Ĺ
#	Со
#	ch
c	ο.
c	ο.
#	No

cd ... done

the subject list to make modifying this script st a subset of subjects easier.

```
for id in `seq -w 1 26`; do
   echo "===> Starting processing of $subj"
```

If the brain mask doesn't exist, create it if [! -f anat/\${subj}_T1w_brain_f02.nii.gz]; then echo "Skull-stripped brain not found, using bet with a fractional intensity thresh # Note: This fractional intensity appears to work well for most of the subjects in # Flanker dataset. You may want to change it if you modify this script for your ow bet2 anat/\${subj}_T1w.nii.gz \ anat/\${subj}_T1w_brain_f02.nii.gz -f 0.2

```
opy the design files into the subject directory, and then
nange "sub-08" to the current subject number
../design_run1.fsf .
../design_run2.fsf .
```

```
ote that we are using the | character to delimit the patterns
# instead of the usual / character because there are / characters
# in the pattern.
sed -i '' "s|sub-08|${subj}|g" \
    design run1.fsf
sed -i '' "s|sub-08|${subj}|g" \
    design_run2.fsf
```

```
# Now everything is set up to run feat
echo "===> Starting feat for run 1"
feat design_run1.fsf
echo "===> Starting feat for run 2"
feat design run2.fsf
        echo
```

Go back to the directory containing all of the subjects, and repeat the loop

The standard design (anti-)pattern for fMRI preprocessing

- Pick a single software package
 - Usually based on considerations other than performance
- String together the tools from that package into a script to run the preprocessing workflow
 - Generally written by a student or postdoc with little software engineering experience
- Hope it keeps working over time...
- Trust that it does the right thing...

#!/bin/bash

#	Ge	ne	ra	te	
#	to	r	un	j	u
fo	r	id	i	n	•
			bi		

echo cd \$subj

f:	Ĺ
#	Со
#	ch
cp	
c	
#	No
#	in
#	in
se	ed

cd .. done

the subject list to make modifying this script st a subset of subjects easier.

```
seg -w 1 26`; do
      ub-$id"
echo "===> Starting processing of $subj"
```

If the brain mask doesn't exist, create it if [! -f anat/\${subj}_T1w_brain_f02.nii.gz]; then echo "Skull-stripped brain not found, using bet with a fractional intensity thresh # Note: This fractional intensity appears to work well for most of the subjects in # Flanker dataset. You may want to change it if you modify this script for your ow bet2 anat/\${subj}_T1w.nii.gz \ anat/\${subj}_T1w_brain_f02.nii.gz -f 0.2

```
py the design files into the subject directory, and then
nange "sub-08" to the current subject number
./design_run1.fsf .
./design_run2.fsf .
```

```
te that we are using the | character to delimit the patterns
   istead of the usual / character because there are / characters
   the pattern.
   -i '' "s|sub-08|${subj}|g" \
   design run1.fsf
sed -i '' "s|sub-08|${subj}|g" \
    design_run2.fsf
```

```
# Now everything is set up to run feat
echo "===> Starting feat for run 1"
feat design_run1.fsf
echo "===> Starting feat for run 2"
feat design run2.fsf
        echo
```

Go back to the directory containing all of the subjects, and repeat the loop

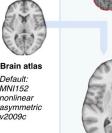
Towards a new design pattern for preprocessing: fMRIPrep

- A robust workflow for preprocessing fMRI data
 - Tested on a large number of fMRI datasets from OpenNeuro
 - Adapt to each dataset and processes it in the best way possible
 - Outputs to BIDS-Derivatives format
 - Provides powerful visualizations :
 - to help identify problems
 - allowing glass-box access to training researchers
 - Currently averaging ~5000 successful runs per week



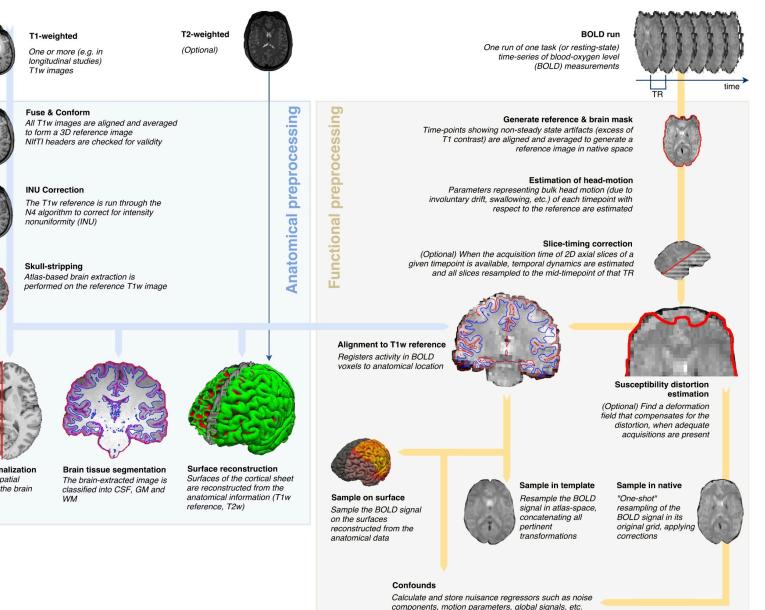






Default: MNI152 nonlinear asymmetr

alignment to the brain



Esteban et al., 2019, *Nature Methods*

RF1MH121867: NiPreps (NeuroImaging PREProcessing toolS)

... the overarching objective of this project is to develop NiPreps, a software framework to perform standardized preprocessing of diverse neuroimaging data.

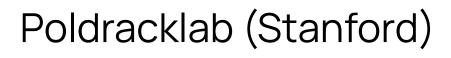
- Aim 1: solidify the foundations of the NiPreps integration.
- Aim 2: enable integrative analysis approaches of heterogeneous data.
- Aim 3: accelerate the dissemination of NiPreps to the neuroscience community through hackathons and "docusprints".

Why?

In developing fMRIPrep, we learned about critical aspects of neuroimaging workflows. NiPreps is envisioned as a generalization of fMRIPrep.

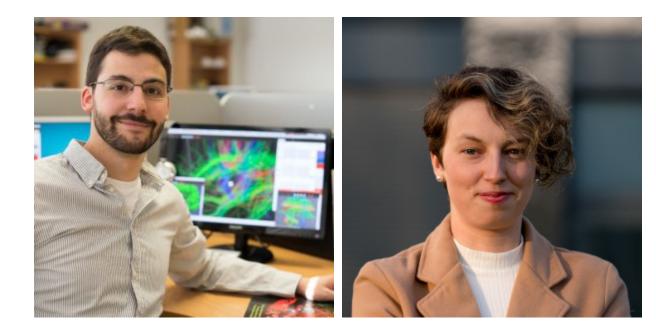
RF1MH121867: Sites







Satterthwaite Lab (Penn)



Esteban Lab (CHUV Lausanne)











Milham Lab (Child Mind Inst)

Ariel Rokem earch Associate Professor n@uw.edu



McKenzie Hagen Graduate student, started 2021



Ezgi Yucel Graduate student (with Ione Fine), started 2017



John Kruper Post-bacc student, started July 2020



Teresa Gomez Post-bacc student, started January 2022



Rokem Lab (UW)

The NiPreps community

>

>

 $\mathbf{\sim}$

>

Welcome

NiPreps

•

Home
NiPreps
BIDS-Apps
Community
Welcome
Membership
New features
Contributing
Code of Conduct
Licensing
Developers

Current members of the GitHub organization

A total of 90 neuroimagers have already joined us. Becoming a member will give you access to additional forums for discussion, subscribing notifications for events and meetings, etc. You can request you are added to the organization by creating a new issue here.



www.nipreps.org



Aim 1: Developing and refining reusable infrastructure/middleware components

- *TemplateFlow*: FAIR Sharing and management of neuroimaging templates and atlases
- *SDCflows*: Integrating susceptibility distortion correction (SDC)
- *NiReports*: A modular visual reports system
- *NiTransforms*: Spatial transforms integration





fMRIPrep Functional MRI "Prep"

The first NiPrep for functional MRI data



sMRIPrep Structural MRI "Prep"

The *NiPrep* for structural (T1-weighted, and T2w) MRI scans, typically used in morphometry analysi and acquired for spatial reference of f/dMRI.



CrowdMRI A database of crowdsourced neuroimaging features

An Internet Service endpoint to collect statistics about data quality, extensible to any crowdsourcing application



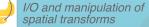
TemplateFlow

ccess, interface, and low across templates

An archive and a client tool to allow pro-grammatic access to neuroimaging templat and atlases, by humans and machines.



NiTransforms



Library supporting the spatial transformatio data formats, and an easy-to-use interface for their application and manipulation.



SDCFlows

Library of tools for estimating and correcting for susceptibility-derived distortions typically affecting EPI images (d/fMRI).

Susceptibility distortion

correction workflows

dMRIPrep Diffusion MRI "Prep

The NiPrep for diffusion MRI scans, typically used in analysis of tractography and white-matter microstructure.



MRIQC

Quality Control (QC) of MRI (structural and functional)

MRIQC produces visual reports for the efficient screening for quality of MRI data, and estimates quality metrics to learn machines flag subpar data.



MRIQCnets

Machine learning for QC and protocol assessmen

Deep Learning models to support higher-level application, such as the deep-MRIQC, the Euler Number Predictor, or the FaceDetector.



Miscellaneous utilities for visualization, reporting, workflow management and fundamental workflows with functionality shared across applications.



NiBabel Access a cacophony of neuroimaging file formats

Library supporting the neuroimaging dat formats (e.g., NIfTI, GIFTI, and CIFTI2).



NiPype Neuroimaging workflows and interfaces in Python

e workflow engine supporting the ecution graph and run time managemen aging tasks, communication, data flow).





The data specification prescribing the formal structure for the neuroimaging data inputs and outputs.

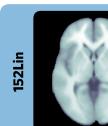
Projects maintained by the

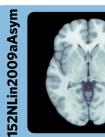
NiPreps community

TemplateFlow: FAIR sharing of neuroimaging templates and atlases

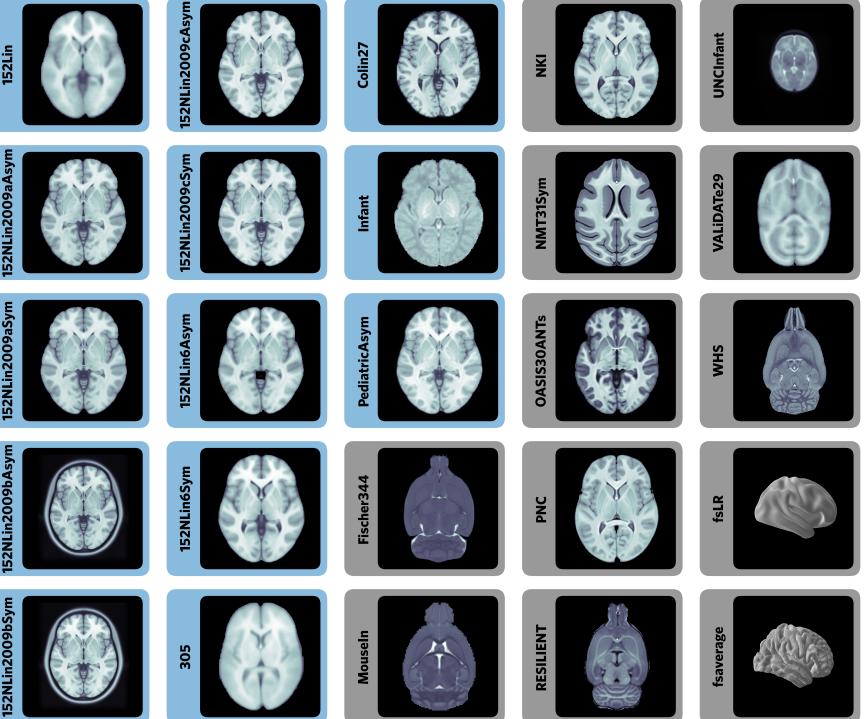
- Templates and atlases are commonly used in neuroimaging research
- There is significant lack of clarity in the use of these templates
 - There are numerous versions of the widely used "MNI template"
- Templateflow provides programmatic access to a database of templates and mappings between them
- Easy to use for humans and machines:

>>> from templateflow import api as tf >>> tf.get("MNI152NLin2009cAsym", suffix="T1w") ...







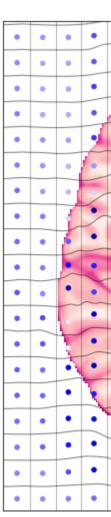


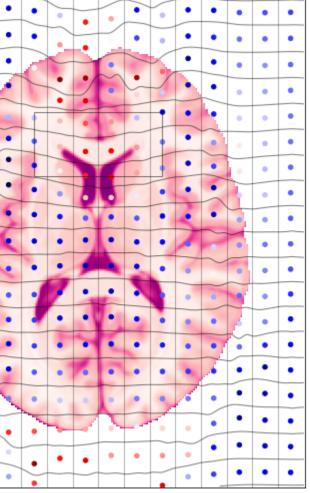


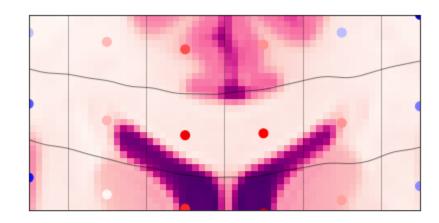
Ciric et al., 2022, *Nature Methods*

SDCFlows: Susceptibility Distortion Correction workflows

- SDCflows aims to provide a unified interface to susceptibility distortion correction methods
 - Defines a shared representation model (B-Spline) for the field map
 - "decouples" estimation and application steps (increasing modularity)
- Overhaul started early 2021 (Esteban et al., OHBM 2021)
 - Faced many technical challenges
 - Requiring numerous bugfixes and "edge" cases
 - Developed new educational materials & Jupyter notebooks



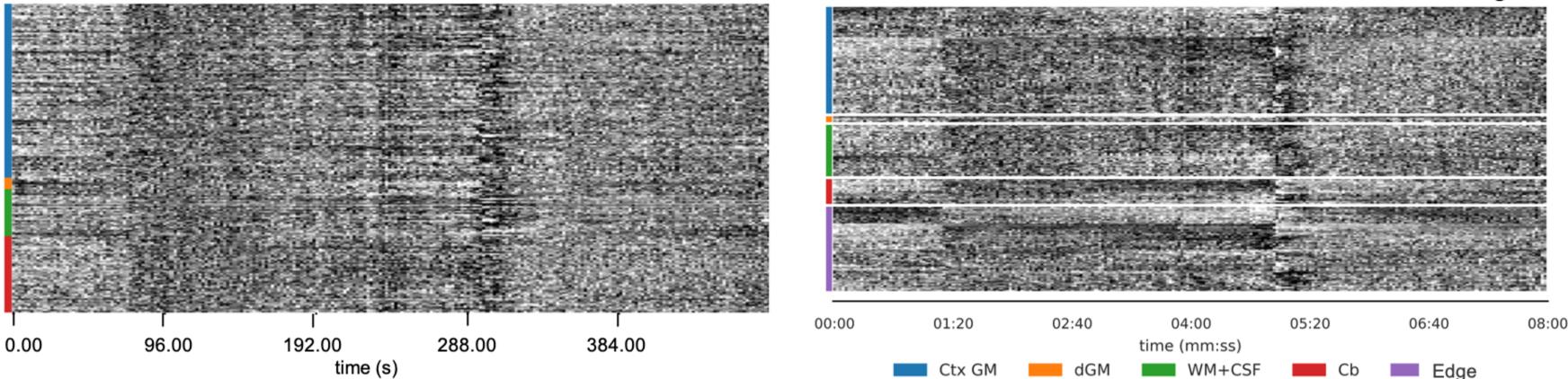




Jupyter Notebook on GitHub: nipreps/sdcflows

NiReports: New visualization tools from MRIQC

Mixed Gambles Task^[21] subject 15 run 02



(Provins et al., ISMRM 2022)

Old Design

- *MRIQC* is a quality control workflow for structural/functional MRI
- Developing a number of visualizations that will go into *NiReports*
 - Added visualization of voxels at the edge of the brain ("crown")
 - Added hierarchical sorting of rows (voxels) to enhance patterns (Aquino et al. 2019)

New Design

Infrastructure: Architectural redesign

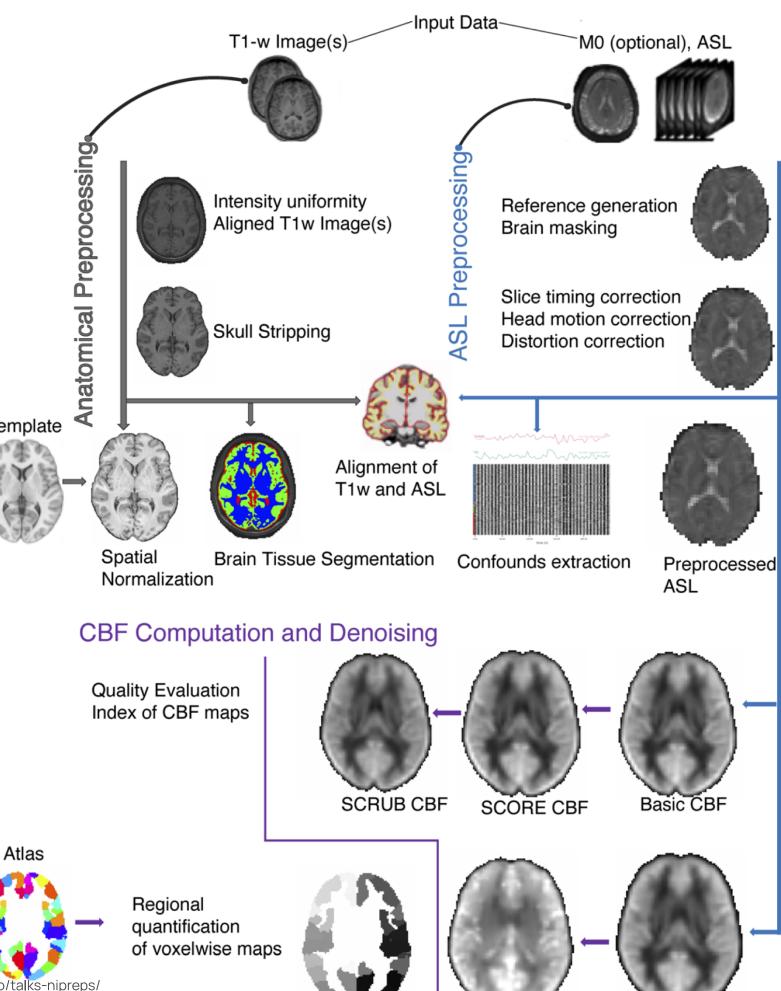
- Problem: fMRIPrep's "one size fits all" design has limitations for emerging use cases Archiving preprocessing results requires balancing storage costs against possible
 - **USE CASES.**
 - Including alternative algorithms requires custom code to integrate.
- Solution: Accept pre-computed derivatives and defer computationally cheap operations
 - E.g., Deep learning segmentations and masks can be accepted, skipping fMRIPrep defaults.
 - Multiple template registrations can be archived, analysts may resample BOLD series to different spaces on demand.
- This approach is implemented in SDCFlows and is being generalized to other components.

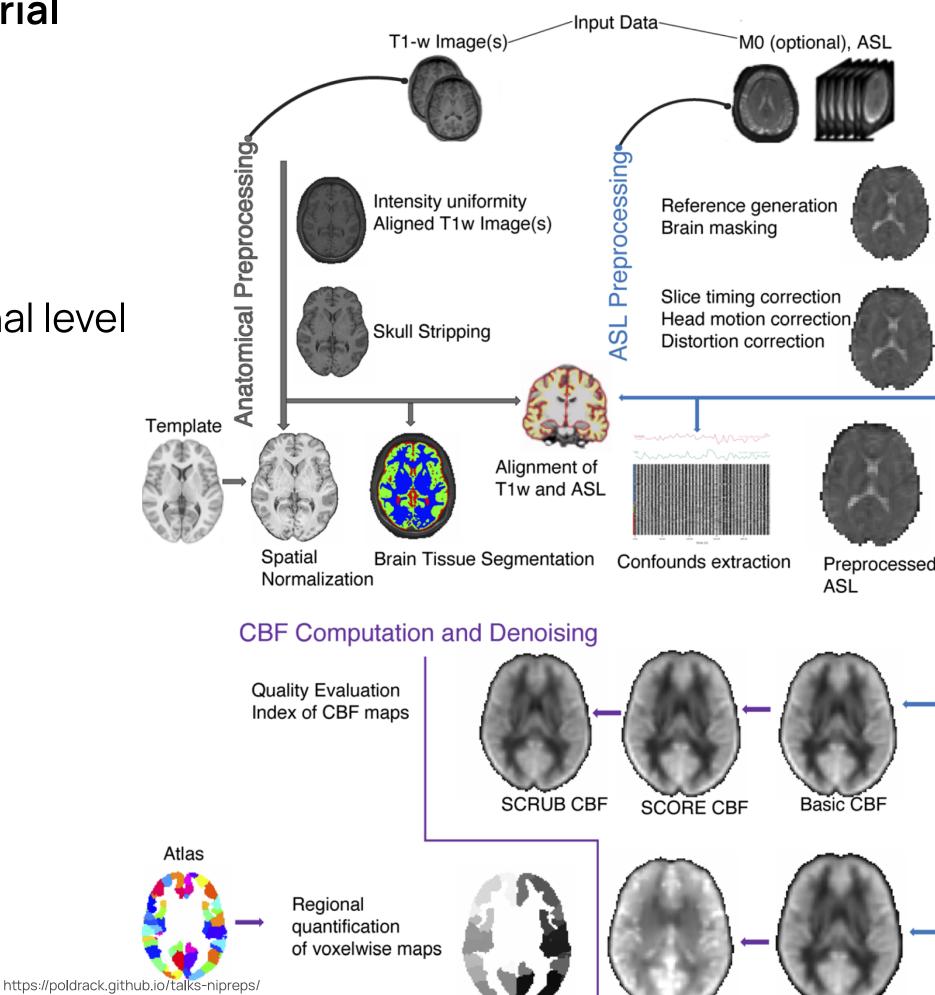
Aim 2: Expand the portfolio of end-user NiPreps

- ASLPrep
- dMRIPrep
- PETPrep
- fMRIPrep-infants (aka *NiBabies*)
- fMRIPrep-rodents (aka *NiRodents*)

Workflows: ASLPrep (cerebral blood flow quantification)

- A robust workflow for preprocessing arterial spin labeling (ASL) data
 - Including cerebral blood flow (CBF) quantification
 - Provides quality evaluation for CBF maps
 - Provides CBF quantification at the regional level using atlases





Workflows: dMRIPrep (diffusion MRI)

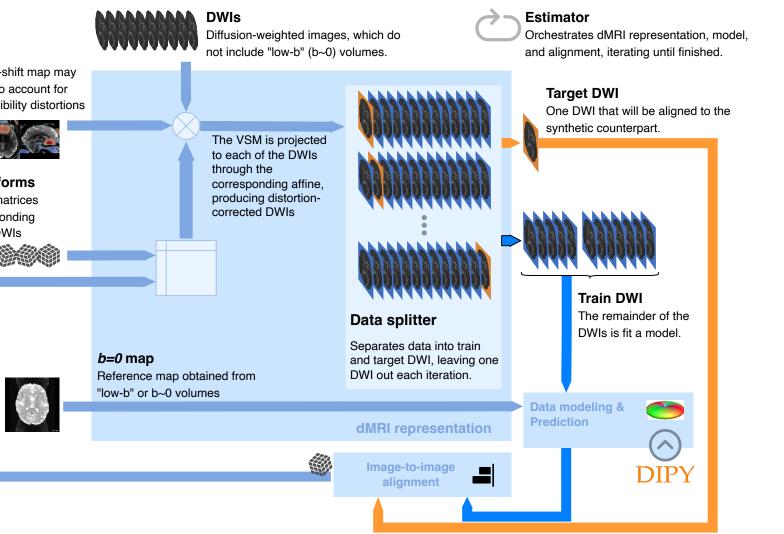
- A workflow for preprocessing of diffusion MRI data
- Development currently focused on eddymotion
 - an algorithm to estimate head-motion (modality-agnostic) and modalityspecific artifacts (eddy currents in the case of dMRI)



A voxel-shift map may be set to account for susceptibility distortions



Transforms Affine matrices corresponding to the DWIs



Preprint: Pisner et al., 2022.

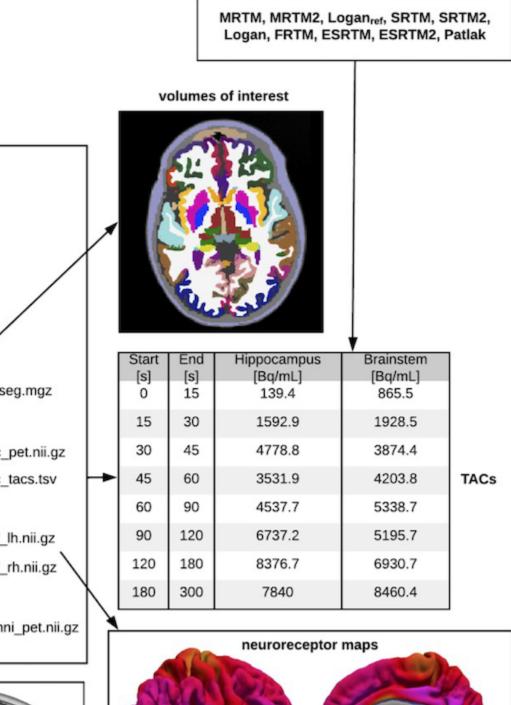
Workflows: PETPrep (positron emission tomography)

- A NiPreps workflow for PET preprocessing
 - Successfully merged petsurfer into nipype (1.8.0)
 - Incorporated nipype implementation of a robust head motion correction workflow (petprep_hmc)
 - Developing a BIDS-Derivatives standard for PET derivatives
 - Model-based head-motion correction leveraging *eddymotion* in progress

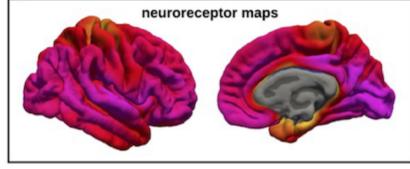
~
study folder
E README
dataset_description.json
participants.tsv
articipants.json
derivatives
🗁 ses-01
🗁 anat 🖊
sub-01_ses-01_seg-gtms
pet
sub-01_ses-01_desc-mc
sub-01_ses-01_desc-mc
surface
sub-01_ses-01_pet_surf_
sub-01_ses-01_pet_surf_
Volume
sub-01_ses-01_space-m
+

neuroreceptor maps

Kinetic model library





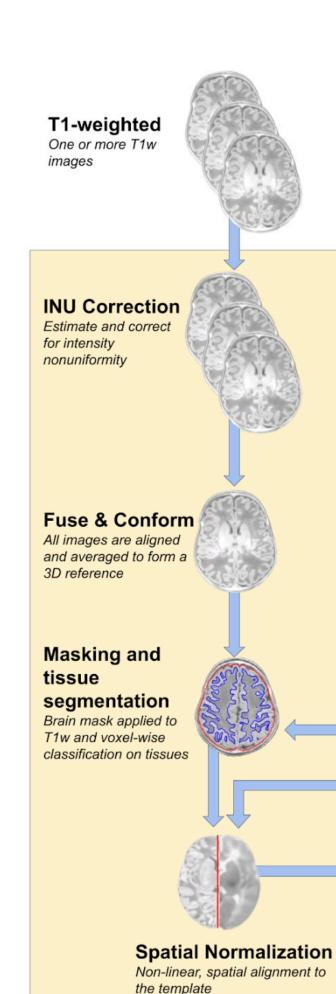


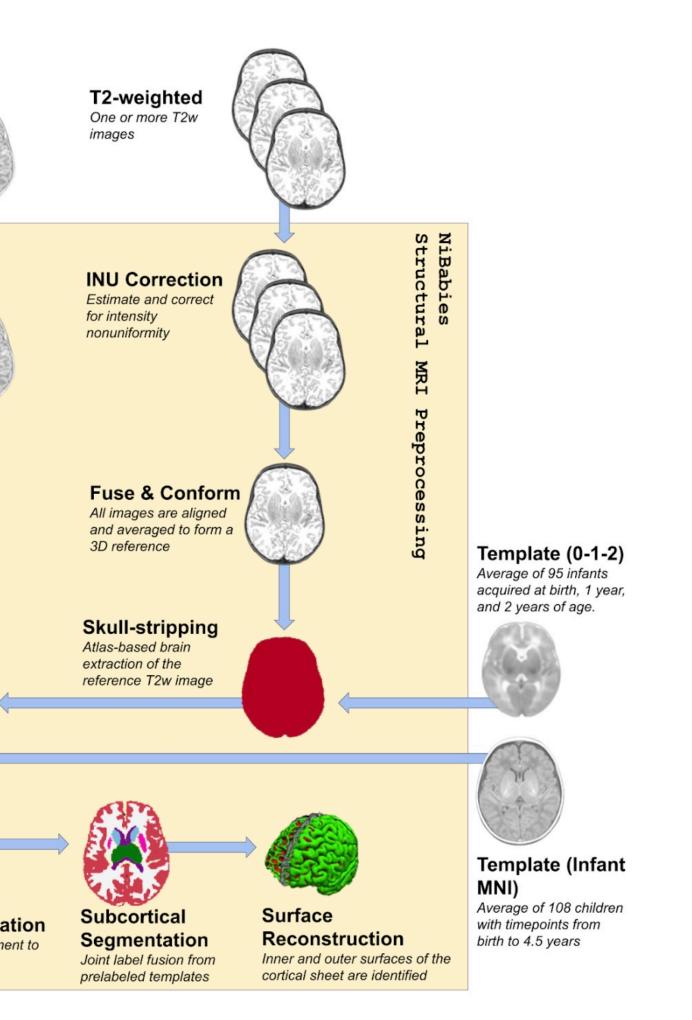
Workflows: fMRIPrep for infants

- Collaboraration with Damien Fair & HBCD team
- New developments
 - Support for pre-computed derivatives (mask, segmentations).
 - Improved robustness and validity of CIFTI-2 outputs.

• Upcoming developments

- Morphometric outputs (cortical thickness, curvature)
- Improvements to susceptibility distortion correction versatility
- T2 assisted surface generation





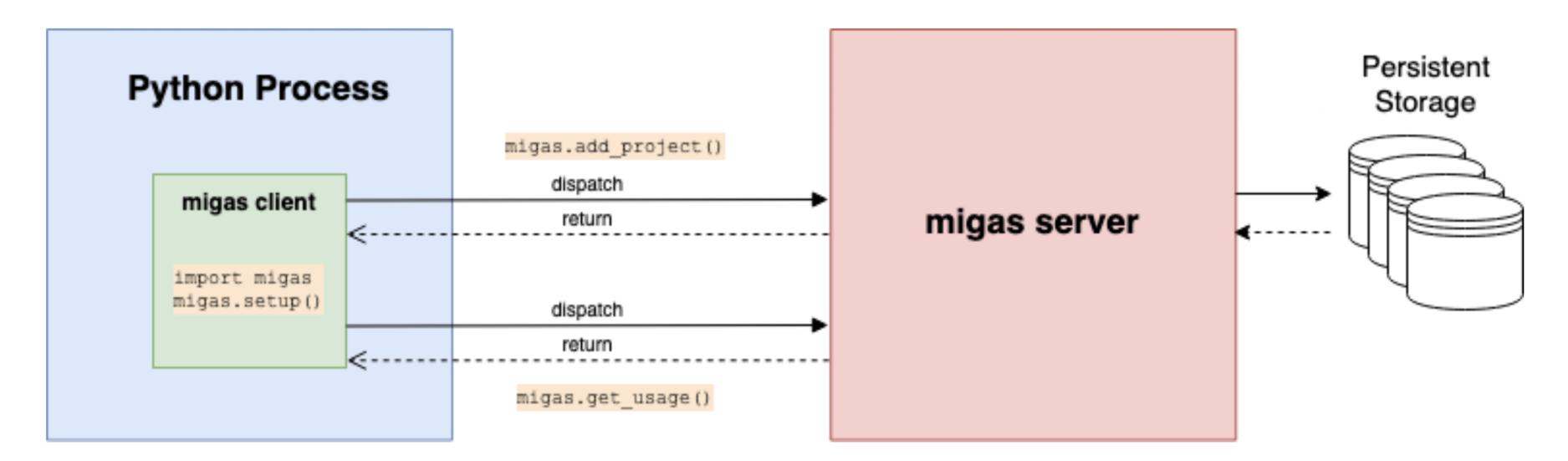
18

Aim 3: Consolidate the NiPreps community

- Project monitoring infrastructure: MIGAS
- Evaluation of cross-workflow reproducibility
- Hackathons and documentation
- Best practices and educational resources
- NMIND: Building common standards for software development

Project monitoring: MIGAS

- An open-source, customizable telemetry solution
- Allows collecting usage information, errors, and status throughout a process's lifetime
- Easy to deploy with various cloud providers (Heroku / GCP / AWS)
- Available as a Python package: https://pypi.org/project/migas/



Reproducibility: Cross-workflow evaluation

- CMI team developed a CPAC implementation of fMRIPrep
 - Able to achieve high levels of reproducibility in connectivity metrics between harmonized workflows
 - Helped identify causes of divergence, such as use of different versions of MNI template

		Matrix Correlation	Edge-wise ICC	I2C2/Discriminability		
Pipeline A	Pipeline B	Schaefer Atlas:		200	600	1000
C-PAC:Default	fMRIPrep-LTS	200 600 1000		0.840	0.742	0.673
e meibelaan	in the city			1.000	1.000	1.000
CC5	fMDIDrop ITC			0.804	0.675	0.602
CCS	fMRIPrep-LTS			1.000	1.000	1.000
000				0.785	0.677	0.620
CCS	C-PAC:Default			1.000	1.000	1.000
				0.732	0.593	0.512
ABCD-BIDS	fMRIPrep-LTS			1.000	0.998	0.995
				0.716	0.557	0.482
ABCD-BIDS	C-PAC:Default			1.000	0.991	0.985
				0.713	0.584	0.509
ABCD-BIDS	DPARSF			1.000	0.999	0.998
				0.642	0.520	0.454
ABCD-BIDS	CCS			0.995	0.989	0.981
				0.657	0.537	0.477
CCS	DPARSF					
				0.990	0.991	0.977
C DAC Default	DRADGE			0.642	0.485	0.420
C-PAC:Default	DPARSF			0.993	0.974	0.958
	001005			0.641	0.495	0.421
fMRIPrep-LTS	DPARSF			0.991	0.985	0.968
	50 5	0.0 0.2 0.4 0.6 0.8 1.0	0.0 0.2 0.4 0.6 0.8 1.0			

Hackathons and documentation

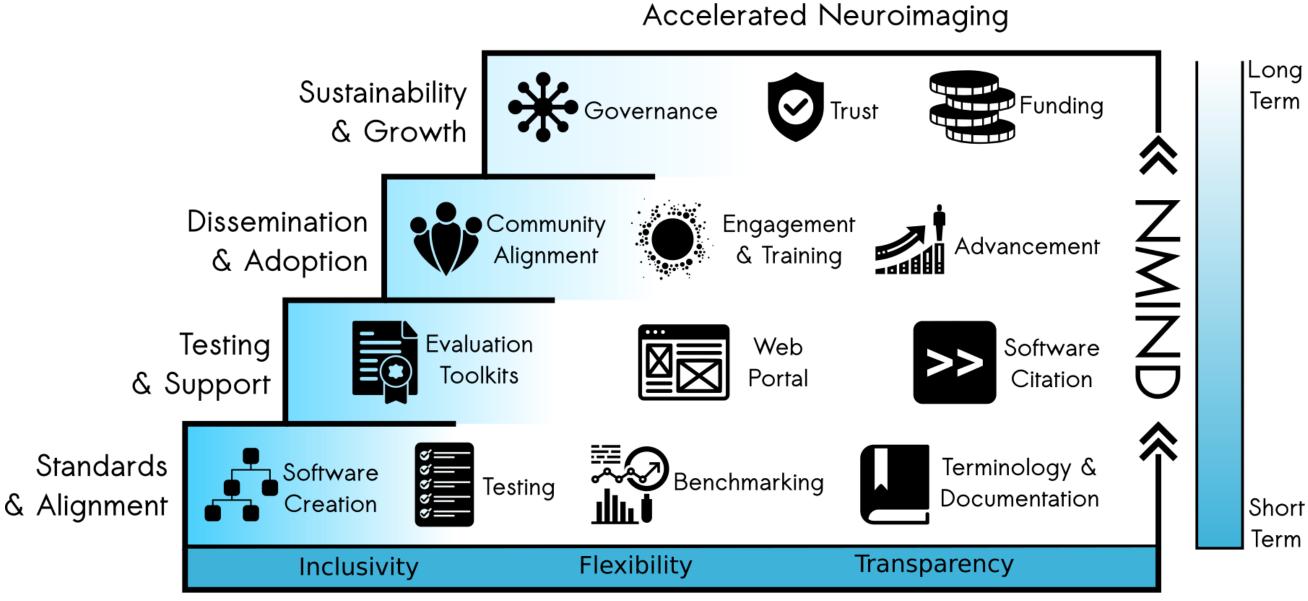
- Held a hackathon/documentation sprint in Glasgow following OHBM 2022
- Will participate in Brainhack Global 2022
- Planning to hold a hackathon/documentation sprint in Montreal in association with OHBM 2023

Best practices and educational resources

- Collaborative QC-Book (educational, ISMRM 2021): https://nipreps.org/qc-book
- MRIQC-SOPs (standard operating procedures)
 - A GitHub template-repository to create and maintain versioned SOPs documentation and checklists.
 - Example: https://nipreps.org/mriqc-sops/
- MRIQC Protocol report (Hagen et al., in preparation)
- Frontiers' research topic on QC of fMRI (Provins et al., under review)
- Biases introduced by defacing in QC (Provins et al., pre-registered report under review)

NMIND: Building common standards for software development

- **NMIND**: Nevermind, this Method Is Not Duplicated
 - Alignment: development and adoption of standards for critical software component
 - Testing: accessible and (semi-)automated mechanisms for evaluating standards compliance
 - Engagement: widespread promotion and adoption of the NMIND collaborative standards



Thank you!