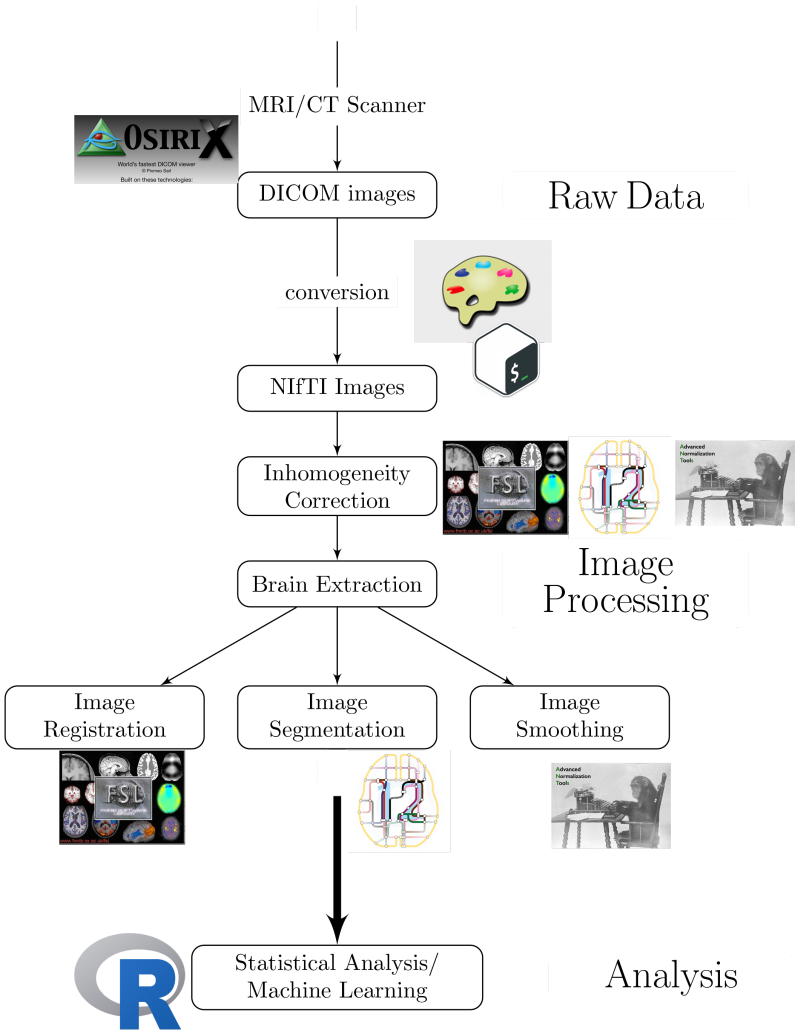
Medical Image Analysis in R

John Muschelli
http://johnmuschelli.com/WCBR_2018.html
Johns Hopkins Bloomberg School of Public Health

Workflow for an Analysis

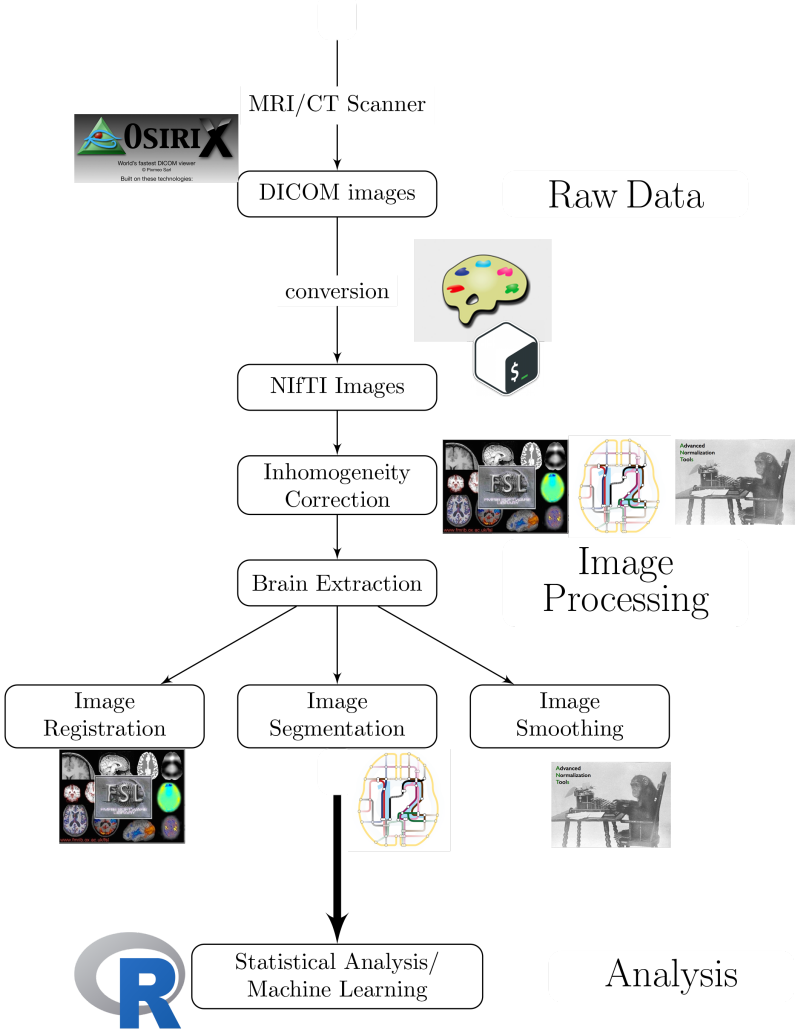
- bash
- FSL
- ANTs
- MRIcroGL
- OsiriX
- SPM 12



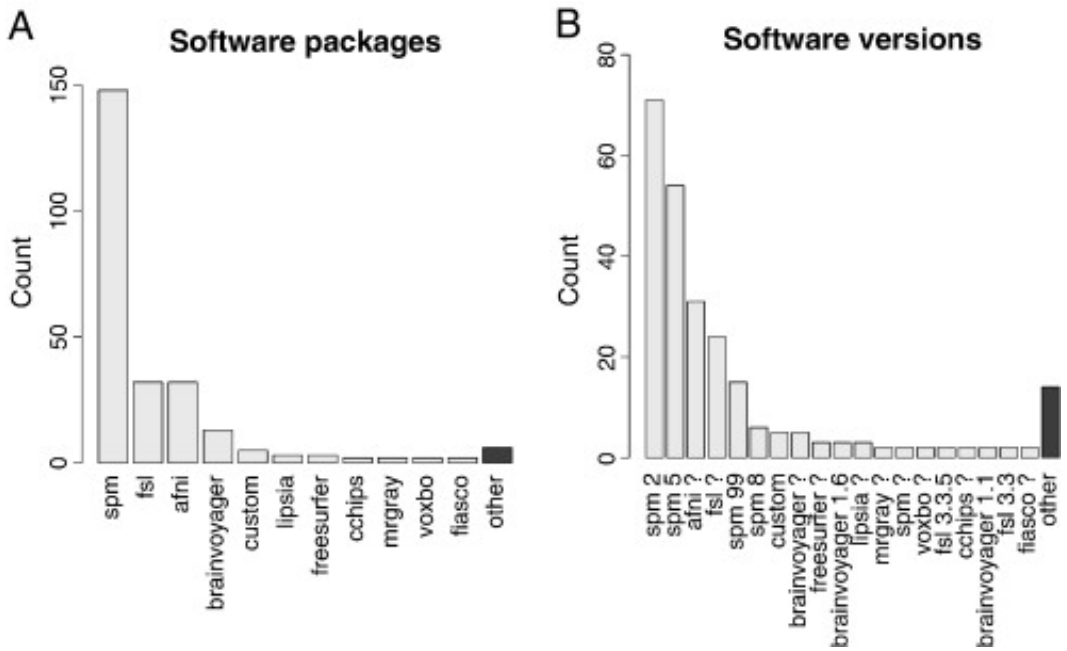
Workflow for an Analysis

Multiple pieces of software used

- all different syntax



It's typical to have lots of software choices



Carp, Joshua. "The secret lives of experiments: methods reporting in the fMRI literature." Neuroimage 63.1 (2012): 289-300.

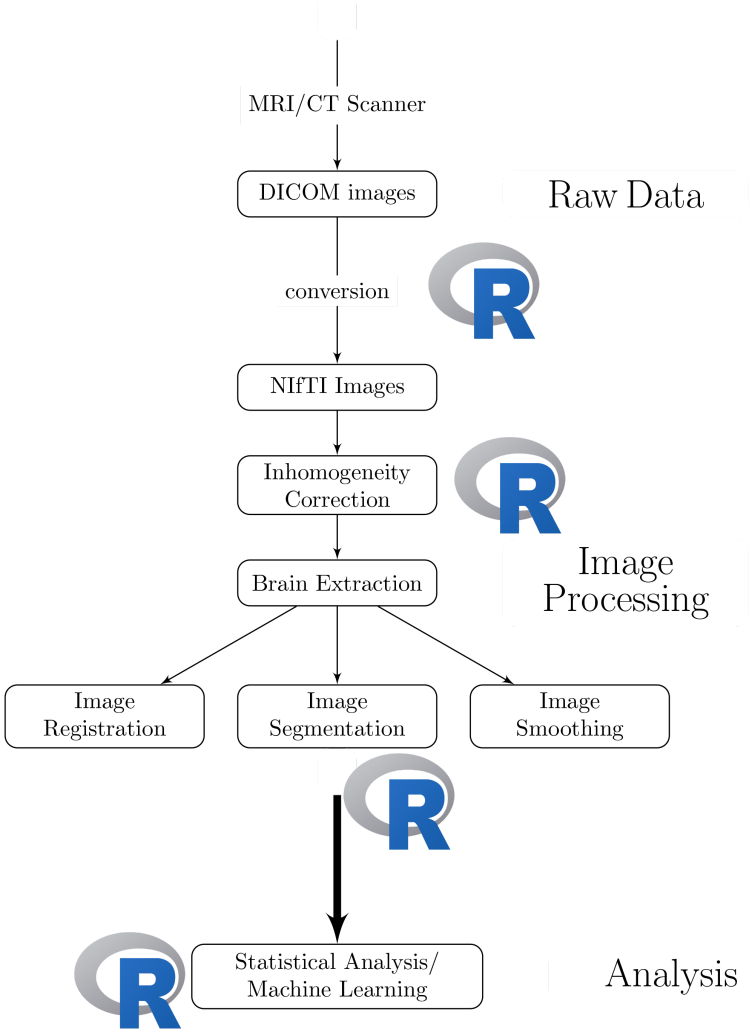
Our Goal:

Lower the bar to entry

- all "one" code (`R`)
  - pipeline tool
  - also "native" `R` code

Complete pipeline

- preprocessing and analysis

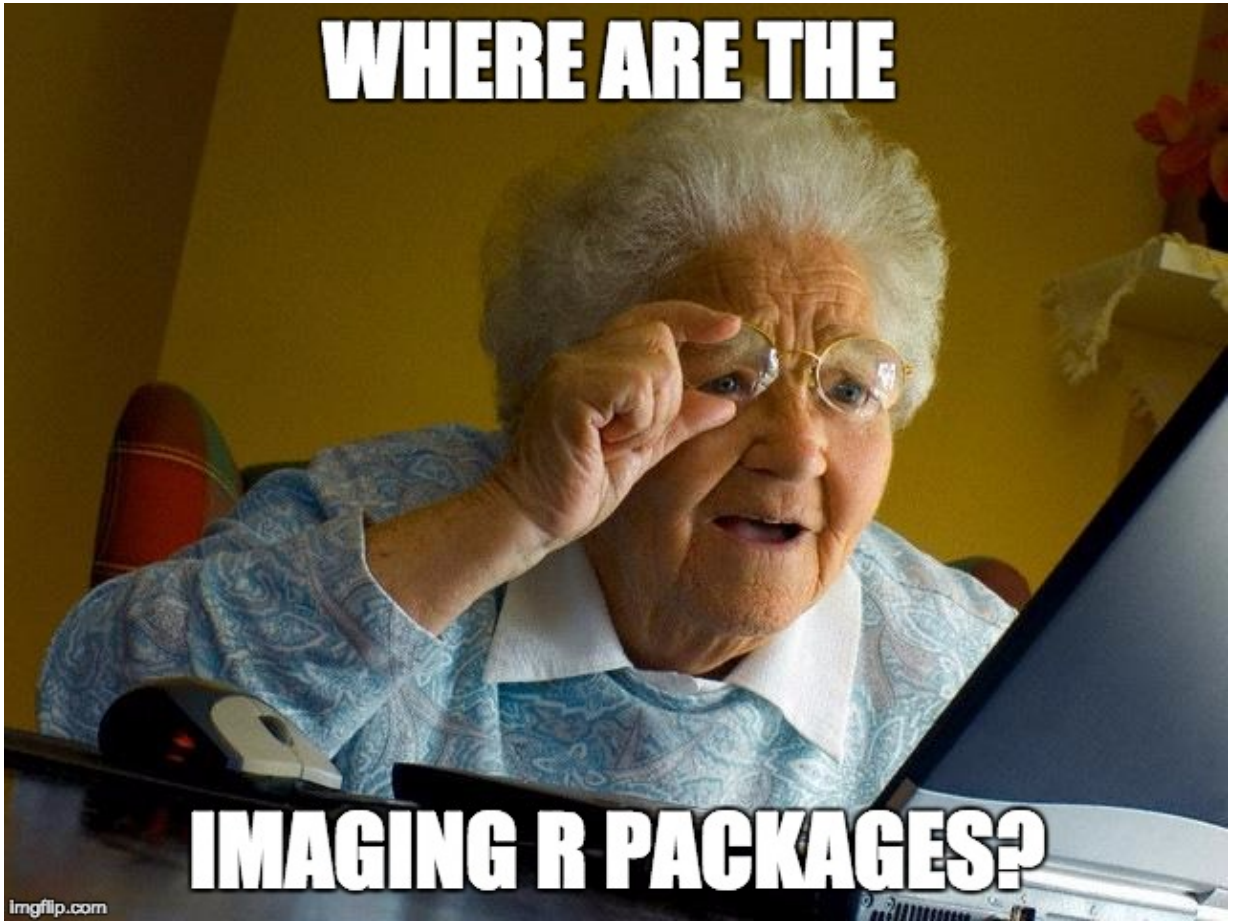


R is a language and environment for statistical computing and graphics.
https://cran.r-project.org/

R is a base set of foundational functions

With (> 11000) user-written packages

What did R have for medical imaging?



https://imgflip.com/memegenerator/Grandma-Finds-The-Internet

What did R have for medical imaging?

CRAN Task View: Medical Image Analysis

Maintainer:	Brandon Whitcher
Contact:	bwhitcher at gmail.com
Version:	2016-12-30
URL:	https://CRAN.R-project.org/view=MedicalImaging

Data Input/Output

DICOM

The industry standard format, for data coming off a clinical imaging device, is DICOM (Digital Imaging and Communications in Medicine). The DICOM "standard" is very broad and very complicated. Roughly speaking each DICOM-compliant file is a collection of fields organized into two four-byte sequences (group,element) that are represented as hexadecimal numbers and form a *tag* . The (group,element) combination announces what type of information is coming next. There is no fixed number of bytes for a DICOM header. The final (group,element) tag should be the "data" tag (7FE0,0010), such that all subsequent information is related to the image(s).

- The packages oro.dicom, divest, fmri and tractor.base (part of the tractor project) provide R functions that read DICOM files and facilitate their conversion to ANALYZE or NIfTI format.

ANALYZE and NIfTI

Although the industry standard for medical imaging data is DICOM, another format has come to be heavily used in the image analysis community. The ANALYZE format was originally developed in conjunction with an image processing system (of the same name) at the Mayo Foundation. An Anlayze (7.5) format image is comprised of two files, the "hdr" and "img" files, that contain information about the acquisition and the acquisition itself, respectively. A more recent adaption of this format is known as NIfTI-1 and is a product of the Data Format Working Group (DFWG) from the Neuroimaging Informatics Technology Initiative (NIfTI). The NIfTI-1 data format is almost identical to the ANALYZE format, but offers a few improvements: merging of the header and image information into one file (.nii), re-organization of the 348-byte fixed header into more relevant categories and the possibility of extending the header information.

- The packages RNifti, AnalyzeFMRI, fmri, tractor.base (part of the tractor project), oro.nifti, and neuroim all provide functions that read/write ANALYZE and NIfTI files.

Magnetic Resonance Imaging (MRI)

Diffusion Tensor Imaging (DTI)

- The R package dti provides structural adaptive smoothing methods for the analysis of diffusion weighted data in the context of the DTI model. Due to its edge preserving properties these smoothing methods are capable of reducing noise without compromizing significant structures (e.g., fibre tracts). The package also provides functions for DTI data processing from input,

Bioinformatics Repository: Bioconductor

- centralized bioinformatics/genomics packages
- large community/number of packages (> 1300)
- published tutorials and workflows
- additional requirements to CRAN (e.g. packages need vignettes)

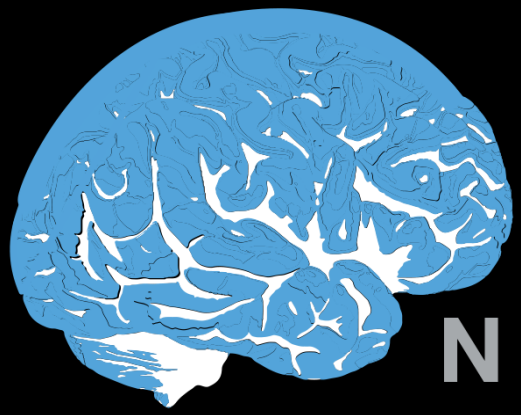
Bioinformatics Repository: Bioconductor

- team of developers/maintainers
- multiple grants of support

Python Framework: NiPy

Nipype:
Neuroimaging in Python
Pipelines and Interfaces

- integrates multiple neuroimaging software
- standardizes the syntax
- provides workflows
- allows the user to take advantage of all of Python
  - pandas, scikit-learn, Jupyter notebooks



NEUROCONDUCTOR

An R Platform for Medical Imaging Analysis

What is Neuroconductor?

1. A community of developers and users of R packages for imaging
2. A website https://neuroconductor.org/.
   - with tutorials and help
3. A team helping developers and users (John, Adi Gherman, Ciprian Crainiceanu, Brian Caffo)
4. A centralized repository of maintained packages

Goal: Centralize the packages (currently 62)

List Packages

View Dependency Graph | View Pending Packages

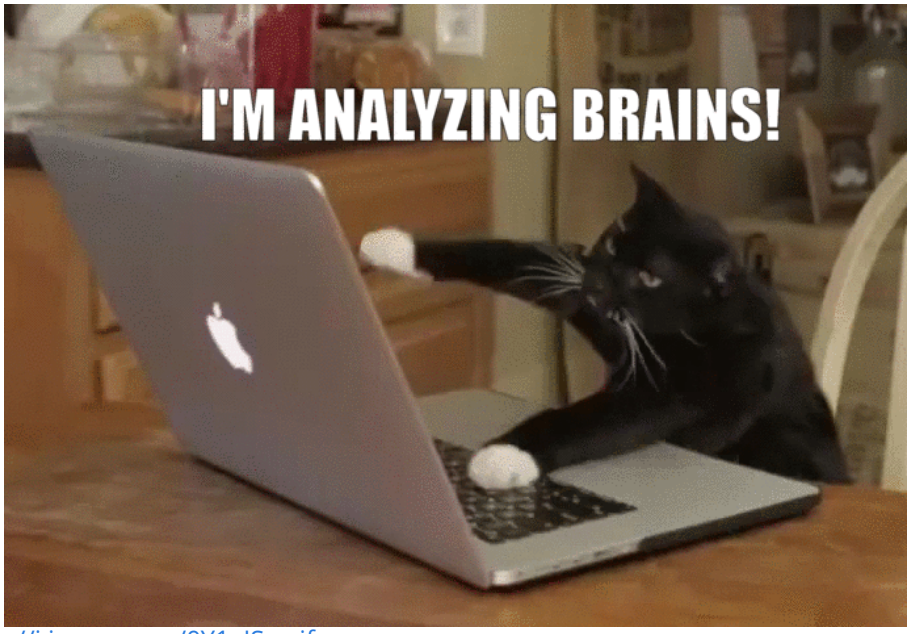
Show 50 entries | Search:

Package Name	Version	Package Title	Maintainer(s)	GitHub URL	Last updated
ANTsR	0.4.0	ANTs in R: quantification tools for biomedical images	Brian B. Avants	stnava/ANTsR	2017-03-18
ANTsRCore	0.0.0	ANTsRCore: core software infrastructure for ANTsR	Brian B. Avants	stnava/ANTsRCore.git	2017-03-18
brainR	1.4.2.1	Helper Functions to Misc3d and rgl Packages for Brain Imaging	John Muschelli	muschellij2/brainR	2017-05-26
cifti	0.4.2	Toolbox for Connectivity Informatics Technology Initiative ('CIFTI') Files	John Muschelli	muschellij2/cifti	2017-05-26
dcemriS4	0.57.1.2	A Package for Image Analysis of DCE-MRI (S4 Implementation)	Brandon, Whitcher	bjw34032/dcemriS4	2017-05-26
dcm2niir	0.5	Conversion of 'DICOM' to 'NIfTI' Imaging Files Through R	John Muschelli	muschellij2/dcm2niir	2017-02-24
divest	0.3.0.1	Get Images Out of DICOM Format Quickly	Jon Clayden	jonclayden/divest	2017-05-25
EveTemplate	0.99.14.2	JHU-MNI-ss (Eve) template	Jean-Philippe Fortin	Jfortin1/EveTemplate	2017-05-26
extrantsr	2.17.2.3	Extra Functions to Build on the ANTsR Package	John Muschelli	muschellij2/extrantsr.git	2017-05-26
freesurfer	1.6.6	Wrapper Functions for 'Freesurfer'	John Muschelli	muschellij2/freesurfer	2017-05-26
fslr	2.12.6	Wrapper Functions for FSL ('FMRIB' Software Library) from Functional MRI of the Brain ('FMRIB')	John Muschelli	muschellij2/fslr	2017-05-26
gifti	0.7	Reads in Neuroimaging 'GIFTI' Files with Geometry Information	John Muschelli	muschellij2/gifti	2016-11-09
ITKR	0.0.1	ITK in R	Brian B. Avants	stnava/ITKR	2017-02-24
itksnapr	2.1.6	Package of ITK-SNAP	John Muschelli	muschellij2/itksnapr	2017-05-26
kirby21.asl	1.5.1	Example ASL Data from the Multi-Modal MRI Reproducibility Resource	John Muschelli	muschellij2/kirby21.asl	2017-05-03

Neuroconductor

Goal:

Detailed **tutorials** on how to actually perform an analysis



From http://i.imgur.com/0Y1xlSa.gifv.

- http://johnmuschelli.com/neuroc

Solution: Build a Bioconductor-like Solution using Current Tools

Git and GitHub

- Git version control system: stores changes of files

- GitHub is an **online** server of repositories
- Distribute packages and install them via `devtools::install_github`

Continuous Integration: Travis and Appveyor

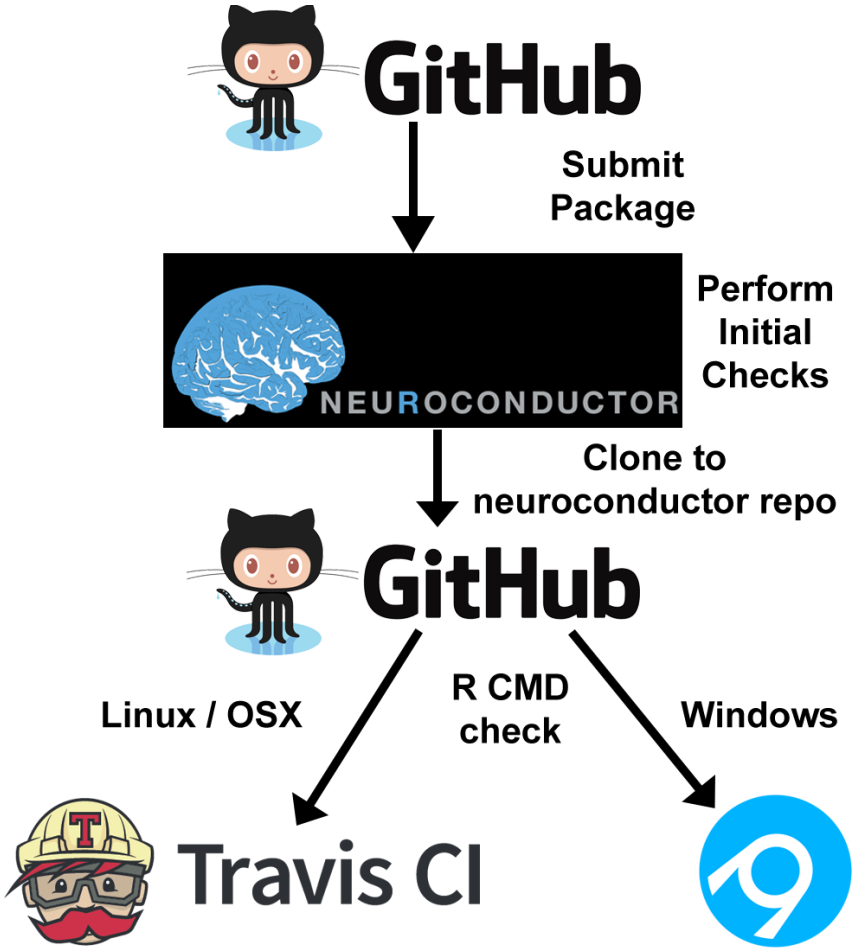
- Builds and checks R packages on Windows (Appveyor) and Linux/OS X (Travis CI)
- Works well with GitHub

Travis CI

Development Pipeline:

Check the package for stability

- check against other imaging software (e.g. FSL)



Helping Developers

- GitHub allows the Neuroconductor team to help fix issues
- Pull Requests to developers
- Standardized checking of Packages (Travis configuration)
- Remove unnecessary hurdles for developers

Image from: https://giphy.com/gifs/medblr-medschool-dr-dres-anatomy-uRb2p09vY8lEs

Benefits of Neuroconductor:

Allow imaging to use all `R` has to offer:

- Statistics and Machine Learning
- Versioning and testing
- Reproducible reports and analyses
- Shiny (web applications)
- Genomics/Imaging analysis in one platform
    - Bioconductor

Neuroconductor Downsides

1. More control over the workflow = more work
2. Users need external software (versions/installation)
3. No control over external software
    - if maintainer changes something, not much recourse
4. Need the content (buy-in from the community)

Questions?

Example Packages

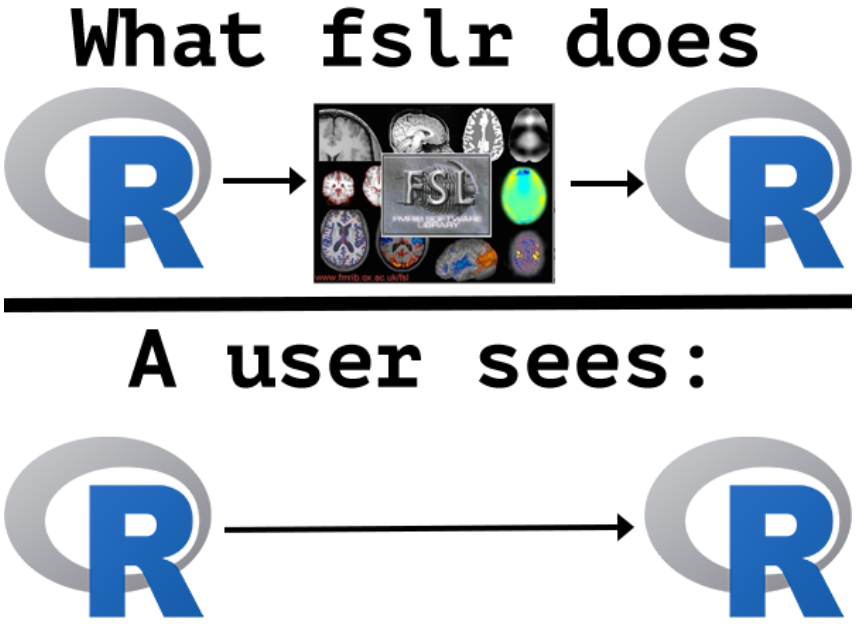
ANTsR

Based on ANTs: Advanced Normalization Tools

- State-of-the-art image processing pipelines
- Group has won challenges for imaging analysis
- Still actively maintained and developed
- Depends on the Insight ToolKit (ITK) medical image processing library

Using R as a Pipeline Tool: fslr

- `fslr` package - call FSL from R
- Requires FSL to be installed (only *nix systems)



spm12r: Wrapper Functions for SPM

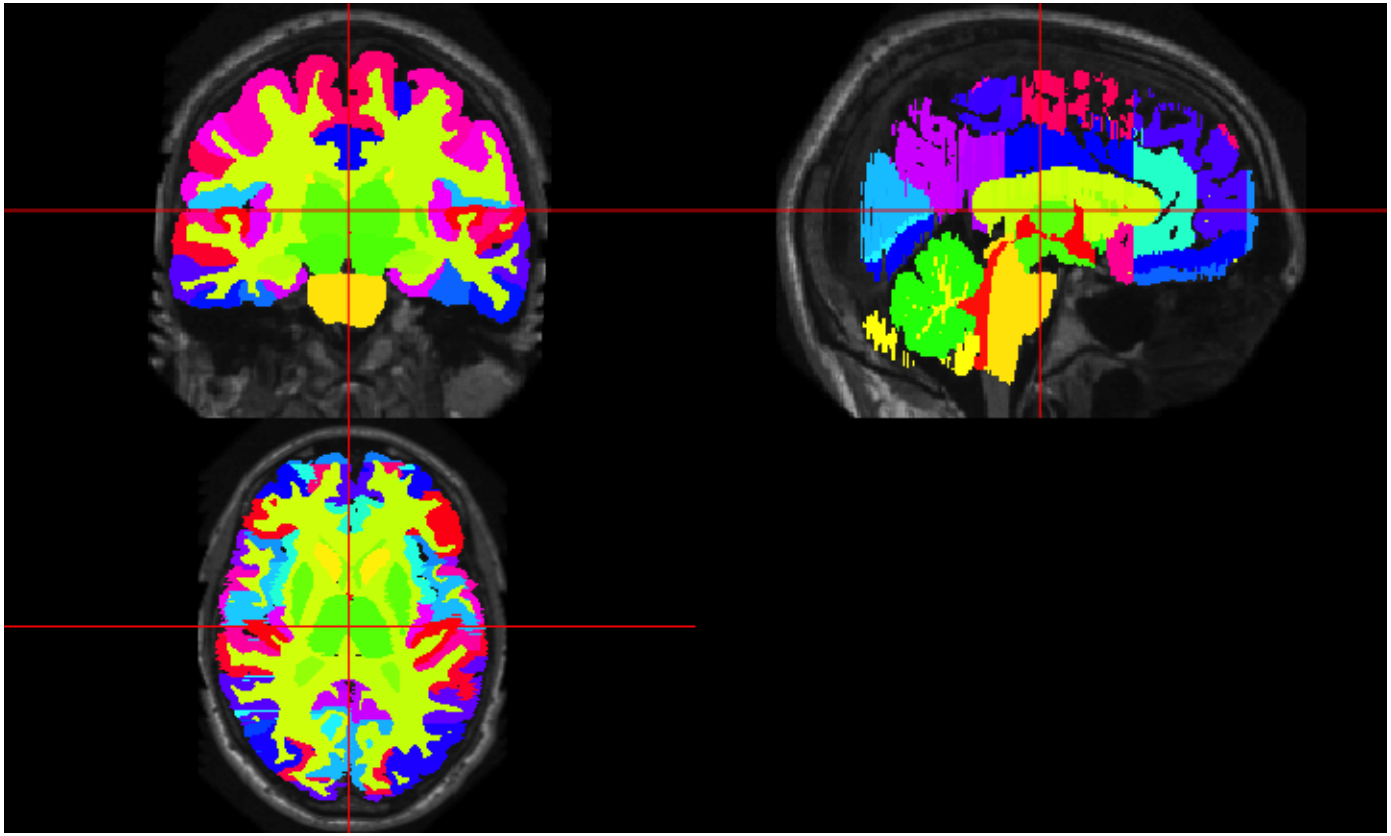
- Wraps some `MATLAB` code to call SPM scripts
- Using `R` syntax (but `MATLAB` runs the code)
- Built from SPM batch commands
- Shown in worked example: http://johnmuschelli.com/talks/fmri_task_processing/index.html#1

neurohcp: Human Connectome Project

- Allows you to download data from Human Connectome Project
- The 1200 Subjects release: behavioral and 3T MR imaging data from 1206 healthy young adult participants. Standardized protocol.
- Tutorial: http://johnmuschelli.com/neuroc/neurohcp

malf.templates: Segmented T1-weighted Images

- Data from the MICCAI 2012 Challenge on Multi-atlas Labelling Data
- From OASIS project and the labeled data as provided by Neuromorphometrics, Inc. (http://Neuromorphometrics.com/)



Current limitations

- R is cross platform, but some packages that depend on *nix system
- Still in beta testing, but more likely to incorporate requests
- Rcpp requires compiled code, (see below)
- Licenses with data can be tricky

Image from: https://imgflip.com/i/22gplr

Training we are providing

Coursera Course: Introduction to Neurohacking In R https://www.coursera.org/learn/neurohacking/

http://johnmuschelli.com/imaging_in_r/ ENAR 2018

Questions?