THE UNIVERSITY OF QUEENSLAND
AUSTRALIA
CREATE CHANGE

Neurodesk

An accessible, flexible, and portable data analysis environment for reproducible neuroimaging

Steffen Bollmann
Senior Research Fellow
School of Information Technology and Electrical Engineering, The University of Queensland, Australia

Acknowledgement of **Country**

The University of Queensland (UQ) acknowledges the Traditional Owners and their custodianship of the lands on which we meet.

We pay our respects to their Ancestors and their descendants, who continue cultural and spiritual connections to Country.

We recognise their valuable contributions to Australian and global society.

Declaration of Potential Conflicts of Interest

I receive research funding from:

1. Oracle for Research - partially fund Neurodesk project via cloud credits
2. Siemens Healthineers

Do we need an accessible, flexible, and portable environment for reproducible Neuroimaging?

The Accessibility problem

Most of our neuroimaging tools require Linux, but are not available in standard package systems

```
user@centos> sudo yum install minc
Loaded plugins: langpacks, ulninfo
No package minc available.
Error: Nothing to do
```

(Users on Debian/Ubuntu are often luckier thanks to Neurodebian)

Then run ccmake .. and set CMAKE_INSTALL_PREFIX to be the desired directory as the above cmake command is ignoring the setting.

make -j 4

This will fail configuring beast.
Edit /home/564/sb1053/minc-toolkit-v2/minc-toolkit-v2/BEaST/CMakeLists.txt
and commend out FIND_PACKAGE(NETCDF) (in two places).

run make -j 4 again.

This will fail to compile /home/564/sb1053/minc-toolkit-v2/minc-toolkit-v2/minctools/progs/mincdump/mincdump.h
Edit this file and replace enum with #define:

The Flexibility problem

Conflicting dependencies -> we often cannot install different versions of software and operating system updates can break our software installations

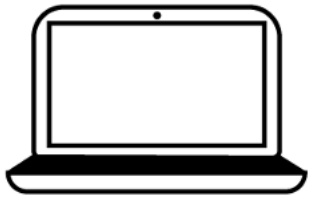
```
freeview.bin: error while loading shared libraries:
 libpng12.so.0: cannot open shared object file: No
such file or directory
```

(for example specific versions of Freesurfer require specific operating system dependencies)

The Portability problem

Reinstalling tools on different platforms is a mess and takes a lot of time …

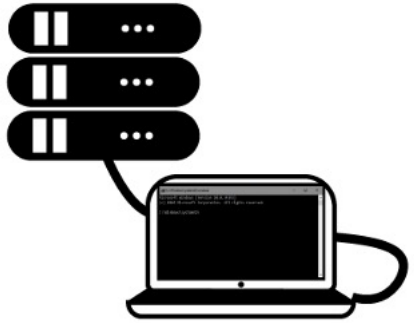
… on your notebook?



… on your lab workstation?

… on a cloud provider?

… on the university's high performance cluster?



The Reproducibility problem

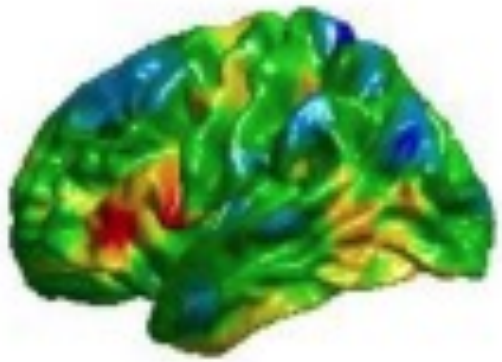
Differing results between software versions ☹

Reproducibility of neuroimaging analyses across operating systems

Tristan Glatard[1,2], Lindsay B. Lewis[1], Rafael Ferreira da Silva[3], Reza Adalat[1], Natacha Beck[1], Claude Lepage[1], Pierre Rioux[1], Marc-Etienne Rousseau[1], Tarek Sherif[1], Ewa Deelman[3], Najmeh Khalili-Mahani[1] and Alan C. Evans[1]

```
expf(1.54051852226257324218750000000)
=4.667009353637695312500 0

expf(1.54051852226257324218750000000)
=4.667009830474853515625 0
```

- glibc 2.5 vs 2.18 deliver different floating-point results
- leads to significant differences in long pipelines



How could we build a data analysis platform that solves these problems by using existing technology and projects?

Acknowledgements

The University Of Queensland AUSTRALIA

Australian Research Data Commons

National Imaging Facility

Korbinian Eckstein
Guillaume Flandin
Martin Grignard
Renzo Huber
Toluwani J. Amos
Lars Kasper
Aina Puce
Satrajit S. Ghosh
Franco Pestilli
Jakub R. Kaczmarzyk
Chris Rorden
Yorguin-Jose Mantilla-Ramos
Megan E. J. Campbell
David J. White
Thomas B Shaw
Benjamin M. Slade
Steffen Bollmann
Kelly G. Garner
Judy D. Zhu
Matthew E. Hughes
Oren Civier
Angela I. Renton
Nigel C. Rogasch
Tom Johnstone
Thanh Thuy Dao
Jeryn Chang
Paul F. Sowman
Paris Lyons
Aswin Narayanan
Levin Kuhlmann
Xincheng Ye
David F. Abbott
Ryan P. Sullivan
Akshaiy Narayanan
Fernanda L. Ribeiro
Jo Morris
Thomas G. Close
Andy Botting
Anthony J. Hannan
Marta I. Garrido
Gary F. Egan
Stefanie Evas
Jason B. Mattingley
Gershon Spitz
Saskia Bollmann
Kexin Lou
Mark Schira
Ashley Stewart

Australian Brain Alliance

Google Summer of Code

ORACLE for Research

OHBM 2020
A Virtual Experience for Engaging Minds & Empowering Brain Science

nectar

Australian Research Container Orchestration Service (ARCOS)

NEURODESK Cite Documentation FAQ Applications Developers Tutorials GitHub News

Quick start

A flexible and scalable data analysis environment for reproducible neuroimaging with Neurodesk.

What is Neurodesk ? | Play | Nectar | Google Colab | Windows | MacOS | Linux

Get started with Neurodesk

Select your setups and follow further instructions in the provided link.

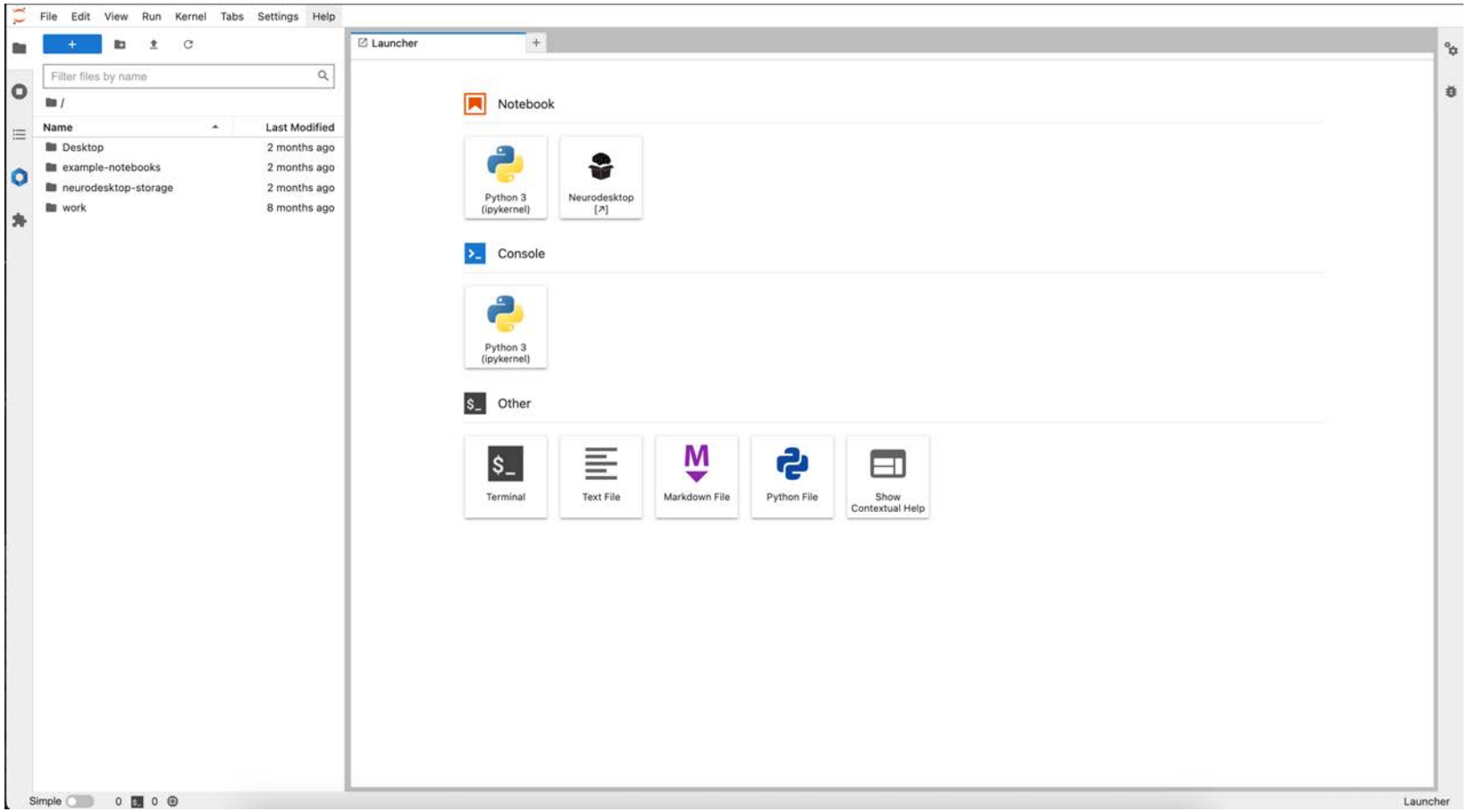
Compute Platform	**Local PC**	HPC	Cloud	Google Colab
Your OS	Linux	**Mac**	Windows	
Interface	**Desktop**	Command Line	Container	VSCode
Processor	**x86**	~~ARM~~	GPU	
Instructions:	Follow the instruction at https://www.neurodesk.org/docs/neurodesktop/getting-started/mac/			

Interactive examples ☺

The next slides are backups

Play – for teaching and quick tests

- jupyter notebook + full desktop
- no authentication required
- no data is saved across sessions

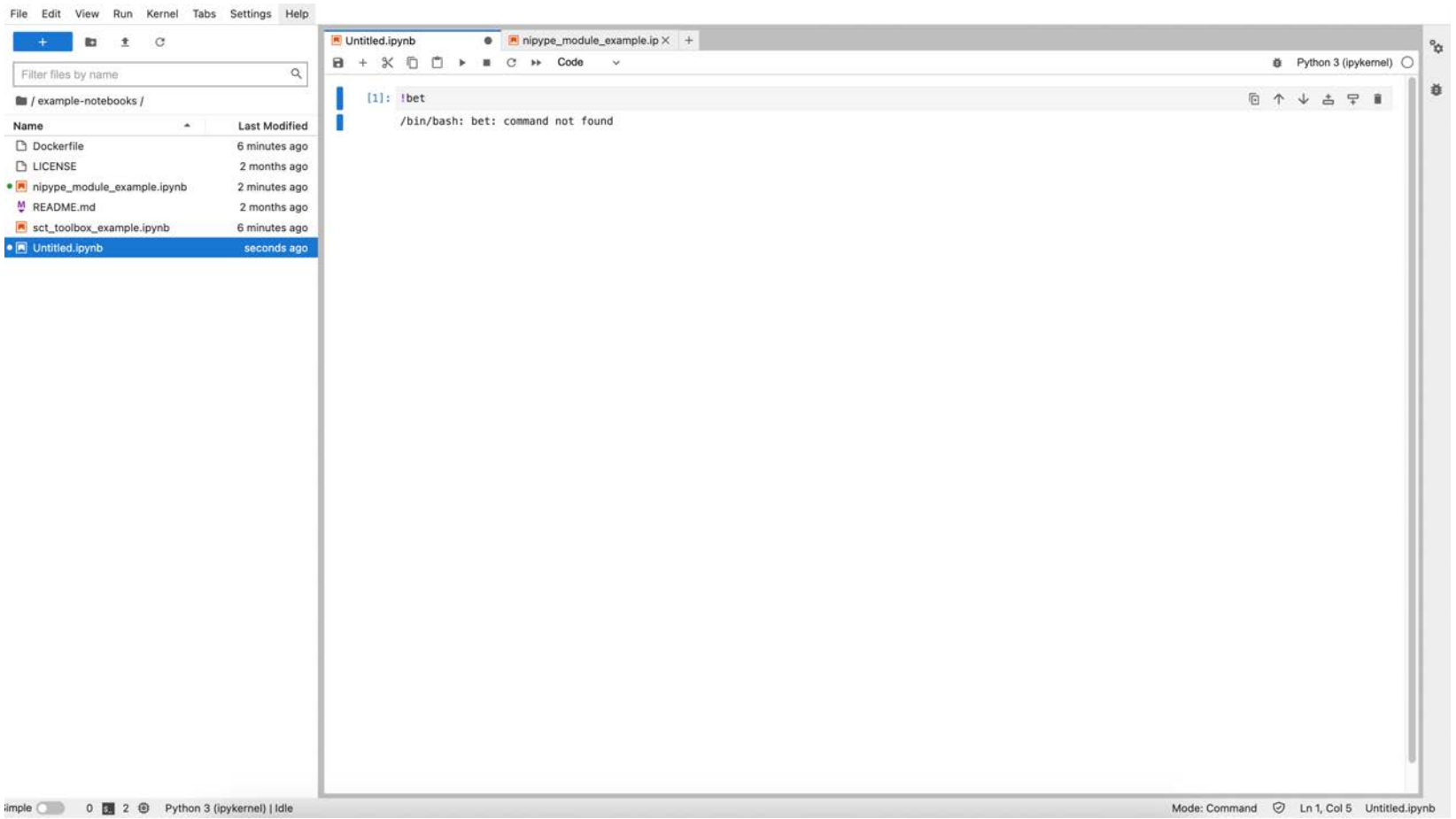


Play – for teaching and quick tests

- full desktop and visual applications
- data can be uploaded with drag and drop
- show bet in desktop and module system there

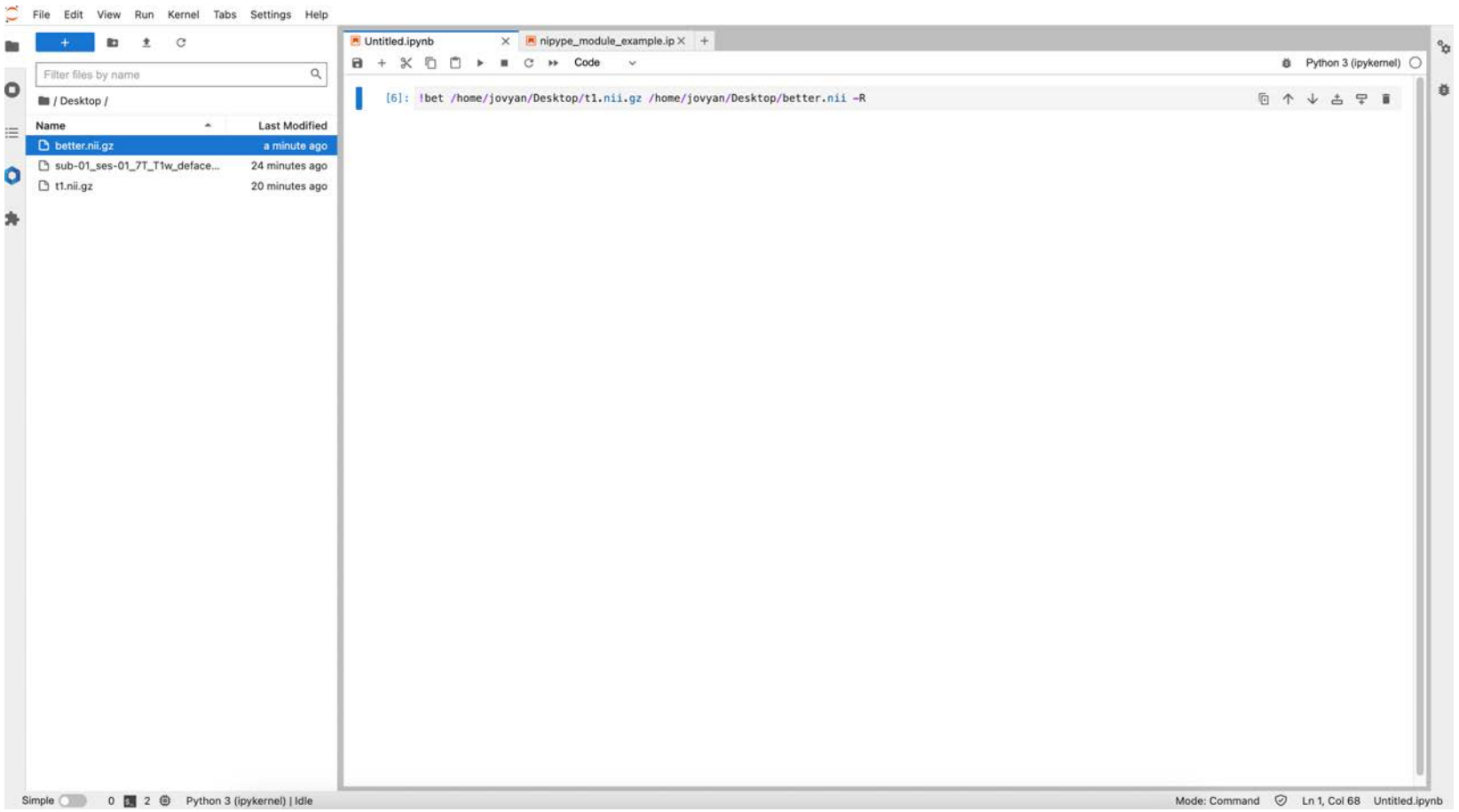
Play – for teaching and quick tests

- jupyter notebook and desktop access the same data and tools!
- 1) show empty notebook without tools loaded



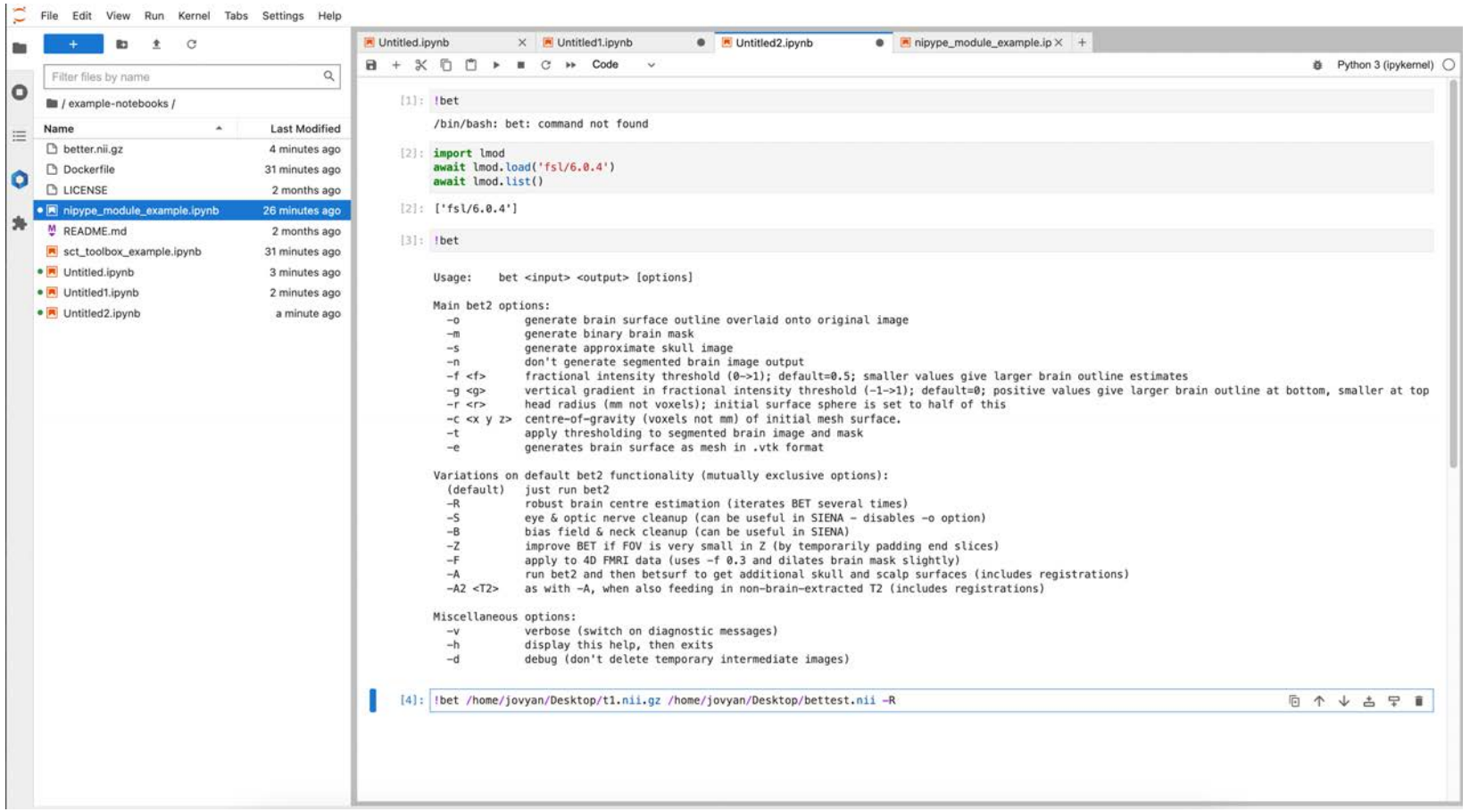
Play – for teaching and quick tests

- show interactive loading of tool and running bet on same T1 data and download data



Play – for teaching and quick tests

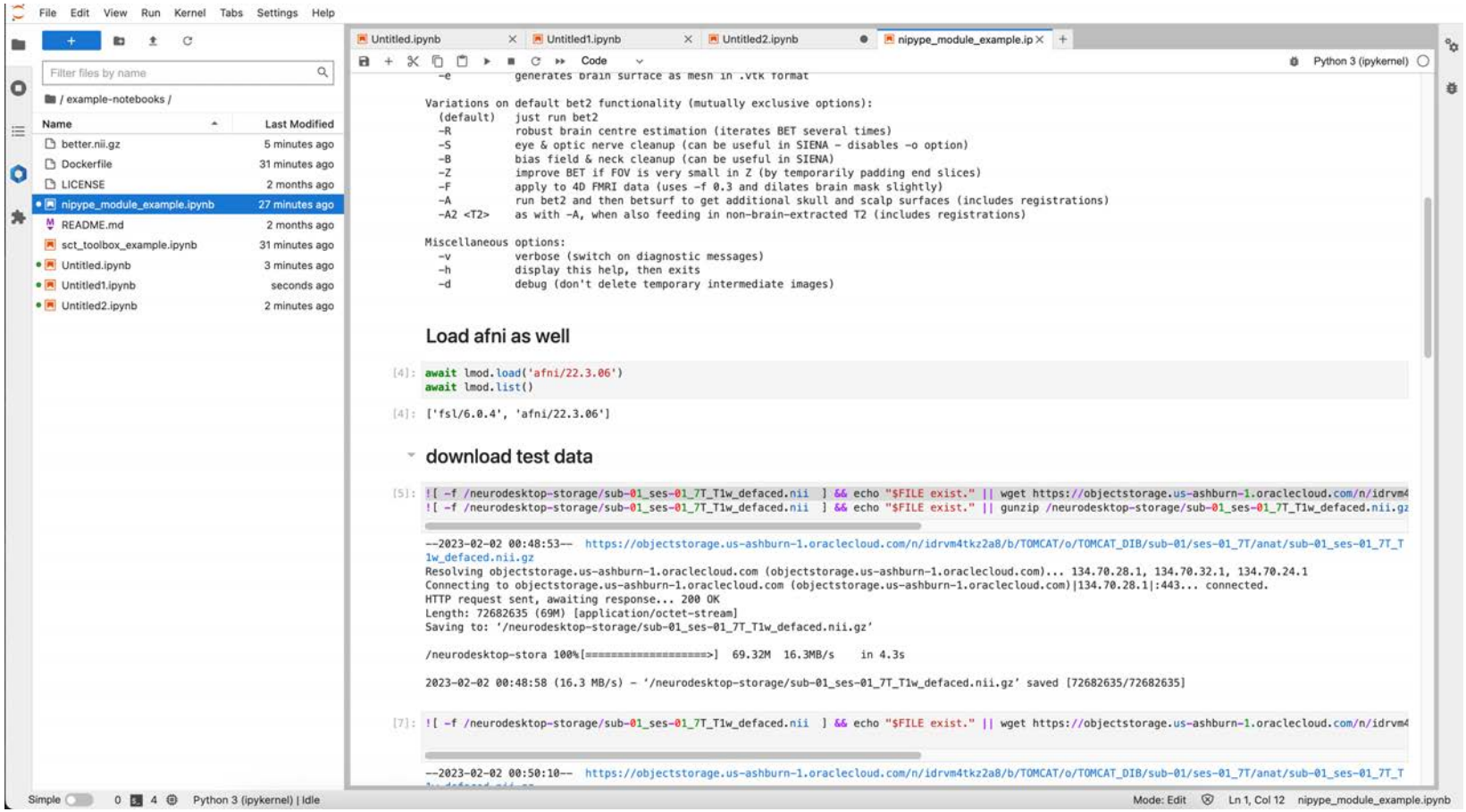
- now, this is not really reproducible, right?
- Let's unload bet and then specify which tool version we are using in code ☺



Play – for teaching and quick tests

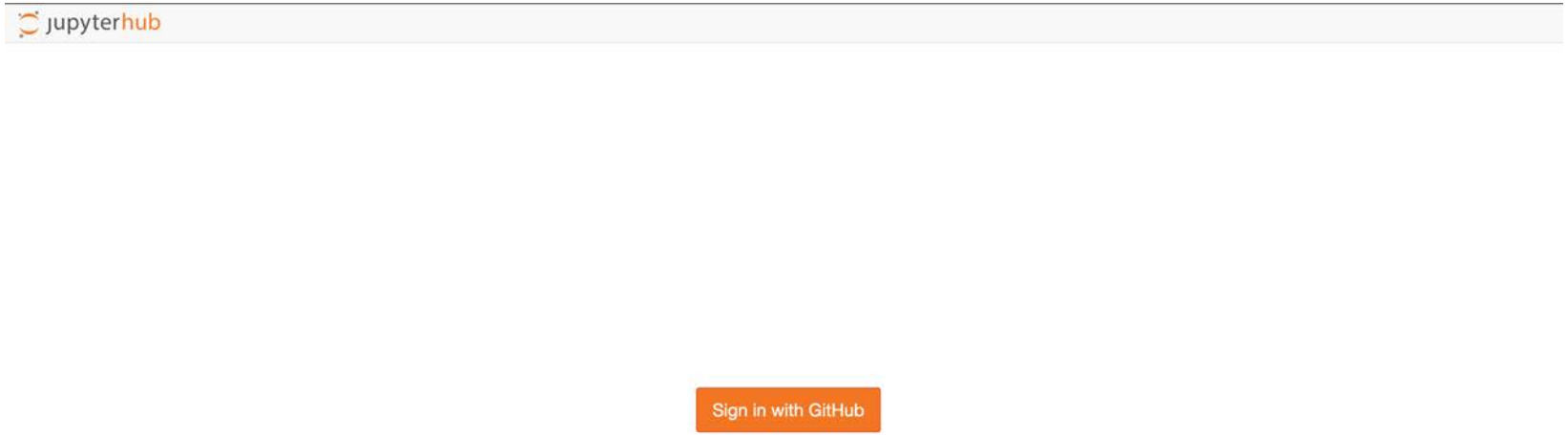
- nipype example

- cool thing is: these notebooks can now be shared and work identically on every neurodesk installation

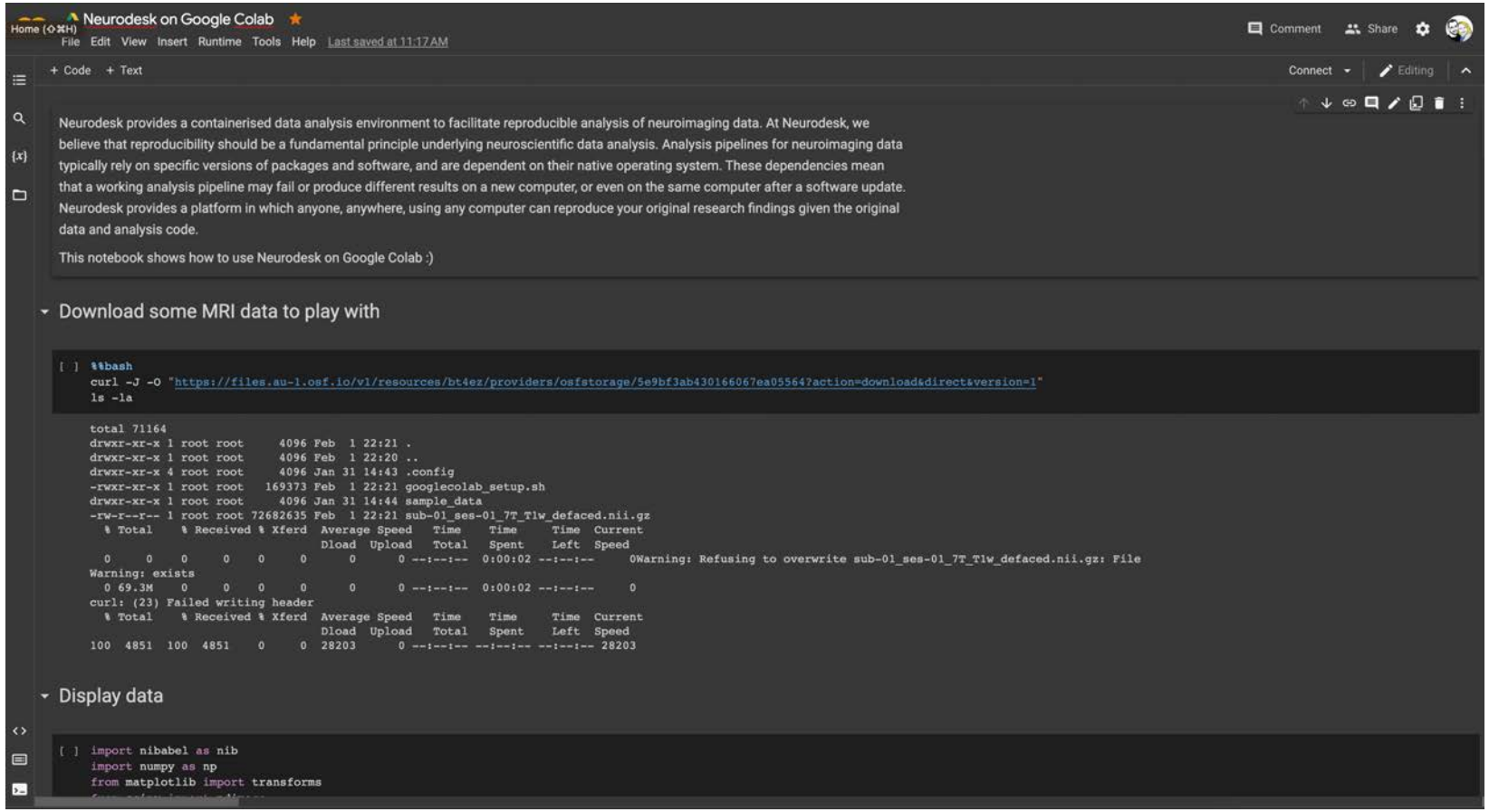


Lab – same thing, but data is saved across sessions

- requires github authentication



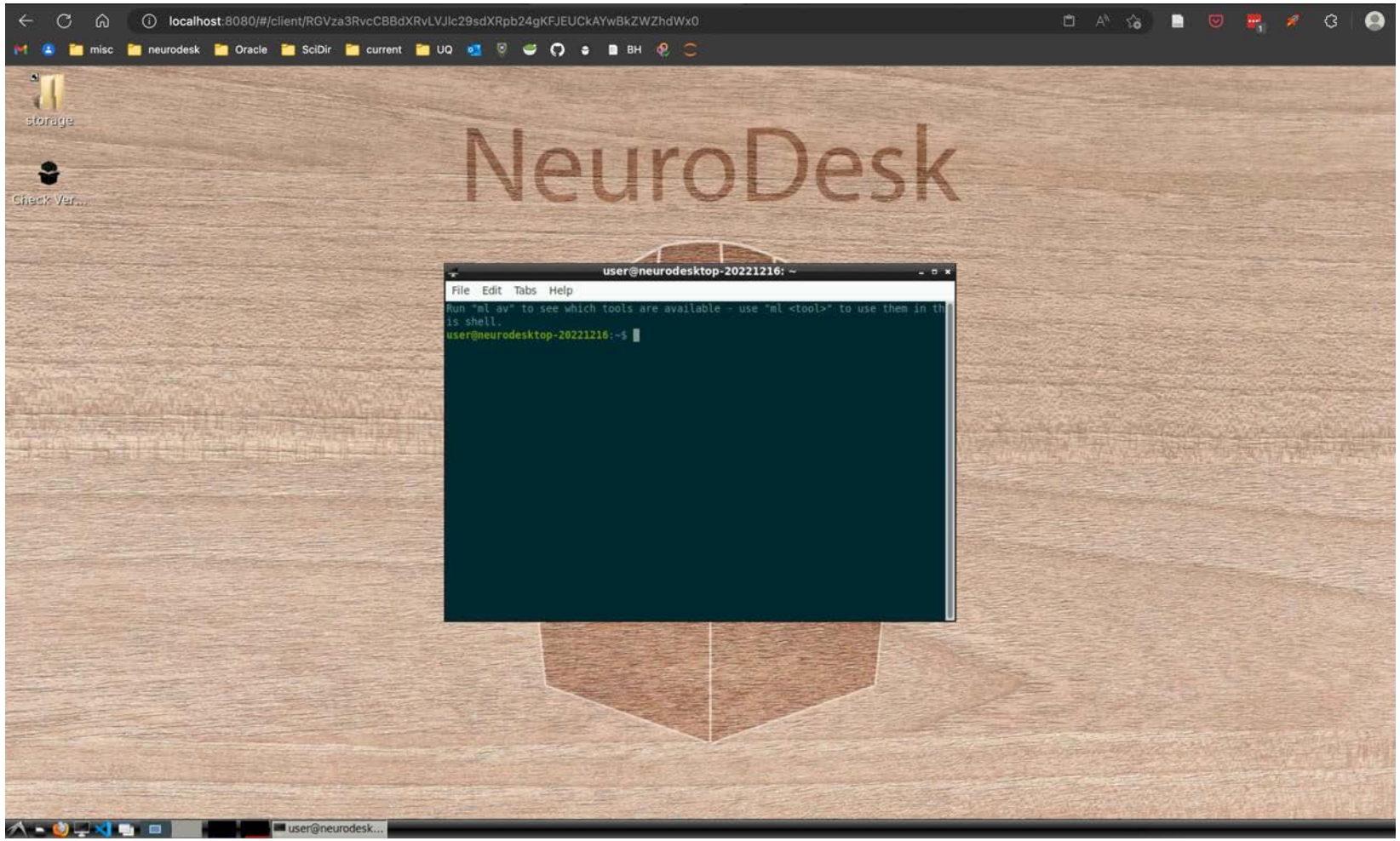
This of course also works on google colab



Running Neurodesk on your own hardware

```
ssh -X -L 8080:127.0.0.1:8080 opc@152.67.98.39
```

```
sudo docker run \
  --shm-size=1gb -it --privileged --name neurodesktop \
  -v ~/neurodesktop-storage:/neurodesktop-storage \
  -e HOST_UID="$(id -u)" -e HOST_GID="$(id -g)"\
  -p 8080:8080 \
  -h neurodesktop-20221216 vnmd/neurodesktop:20221216
```



Running Neurodesk on an HPC

```
ssh uqsbollm@bunya.rcc.uq.edu.au
```

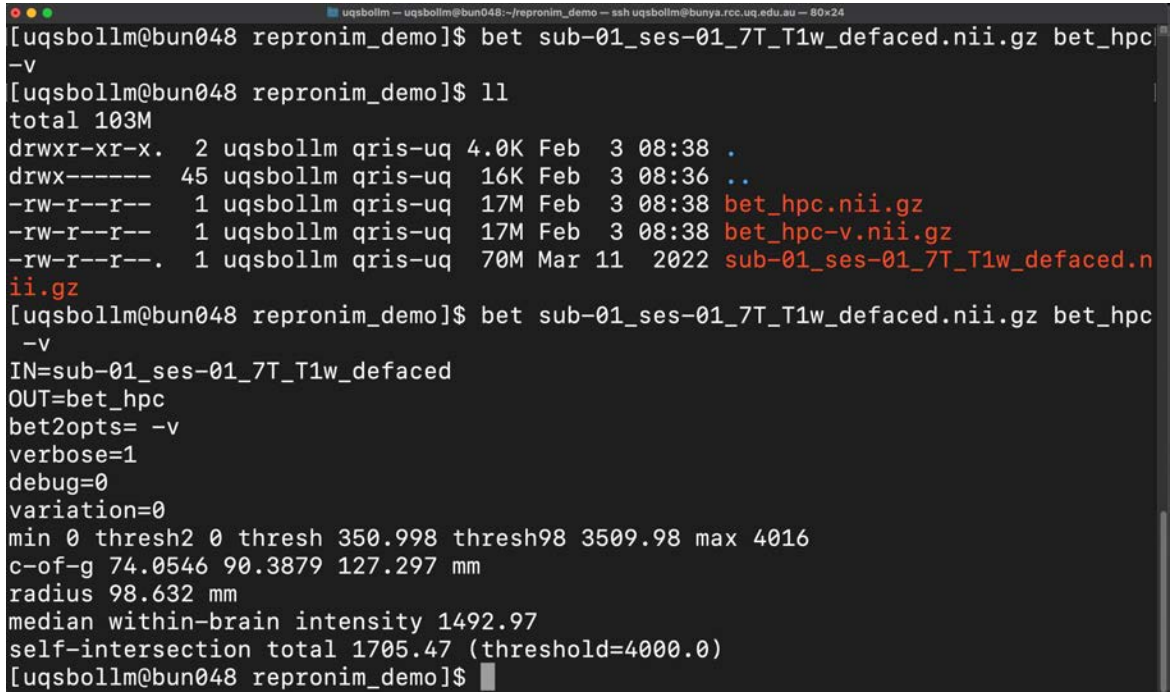
```
salloc --nodes=1 --ntasks-per-node=1 --cpus-per-task=1 --mem=50G --job-name=TEST --time=05:00:00 --partition=general --account=a_barth srun --export=PATH,TERM,HOME,LANG --pty /bin/bash -l
```

```
 module use /scratch/user/uqsbollm/neurocommand/local/containers/modules/
```

```
cd repronimdemo
```

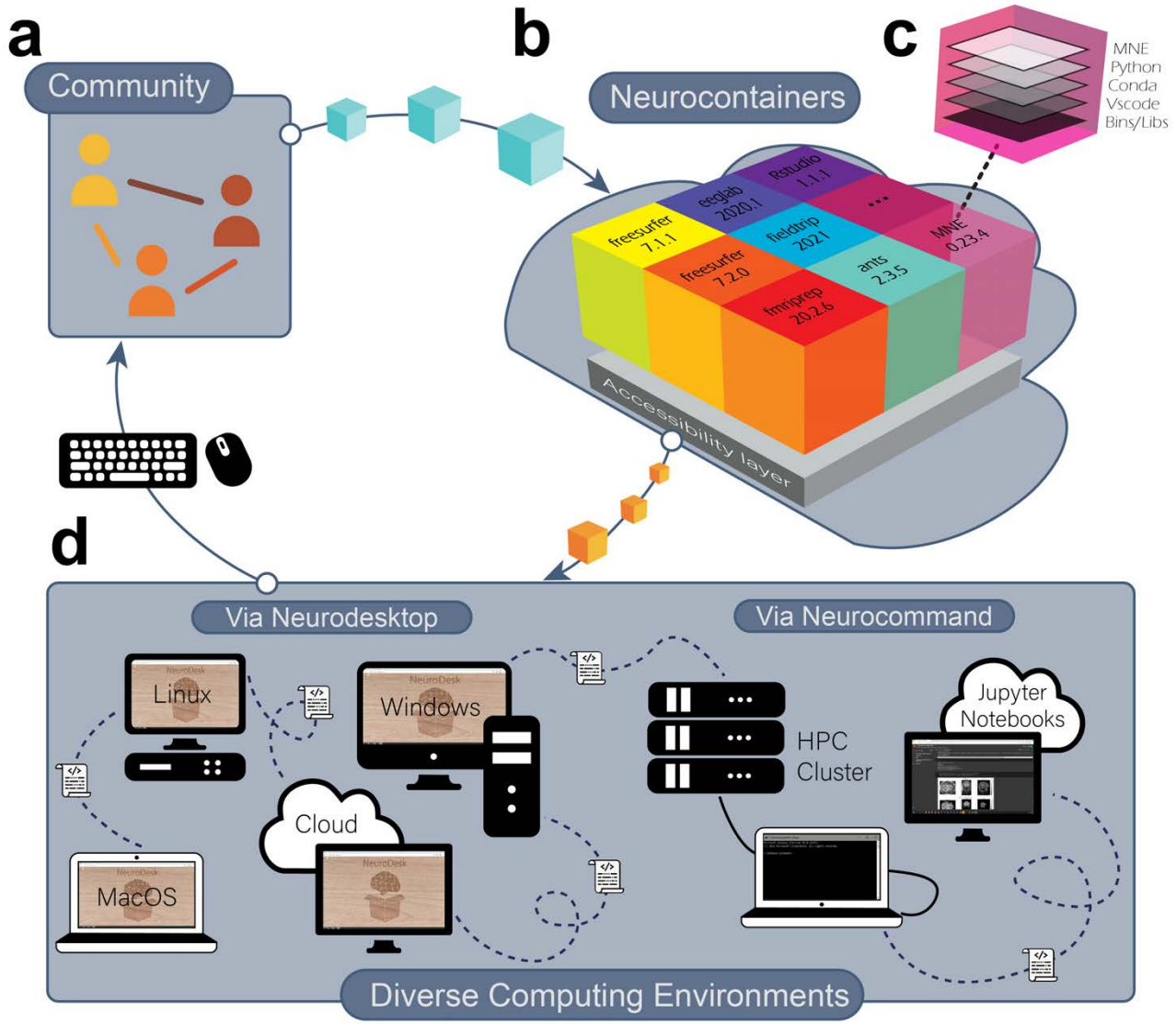
```
ml fsl/6.0.4
```

```
bet sub-01_ses-01_7T_T1w_defaced.nii.gz bet_hpc -v
```



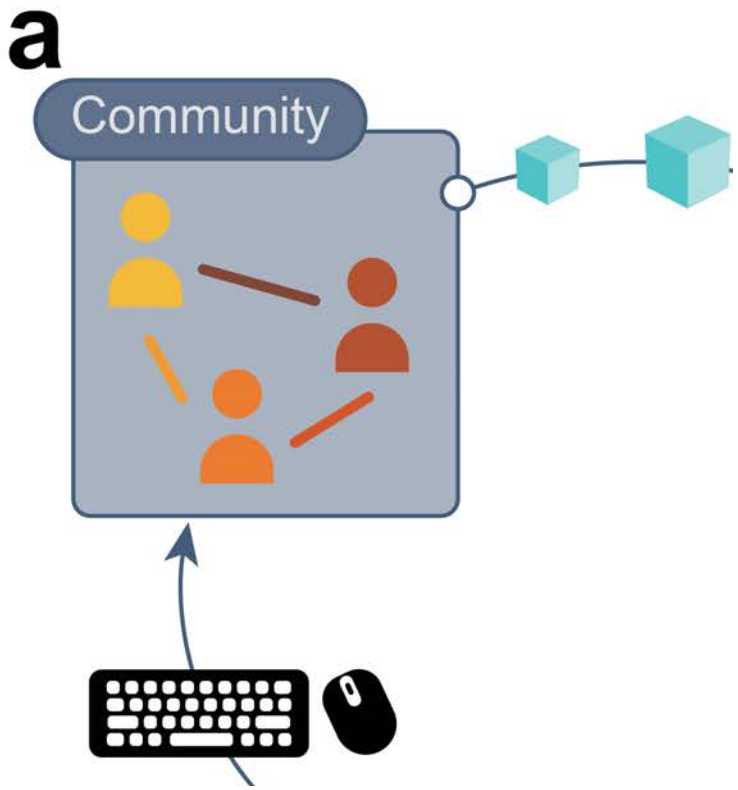
Interactive examples end ☺

The Neurodesk.org Open-Source Platform



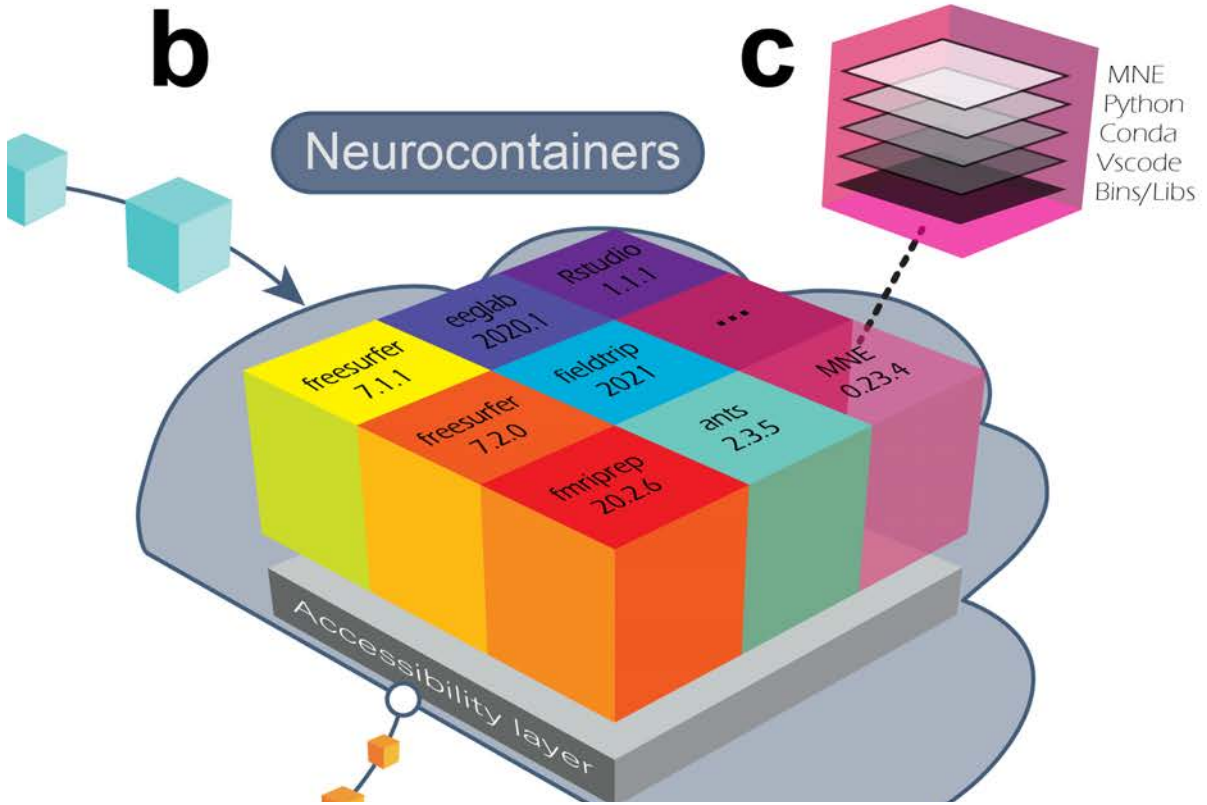
The Neurodesk.org Open-Source Platform

Community builds and maintains software containers



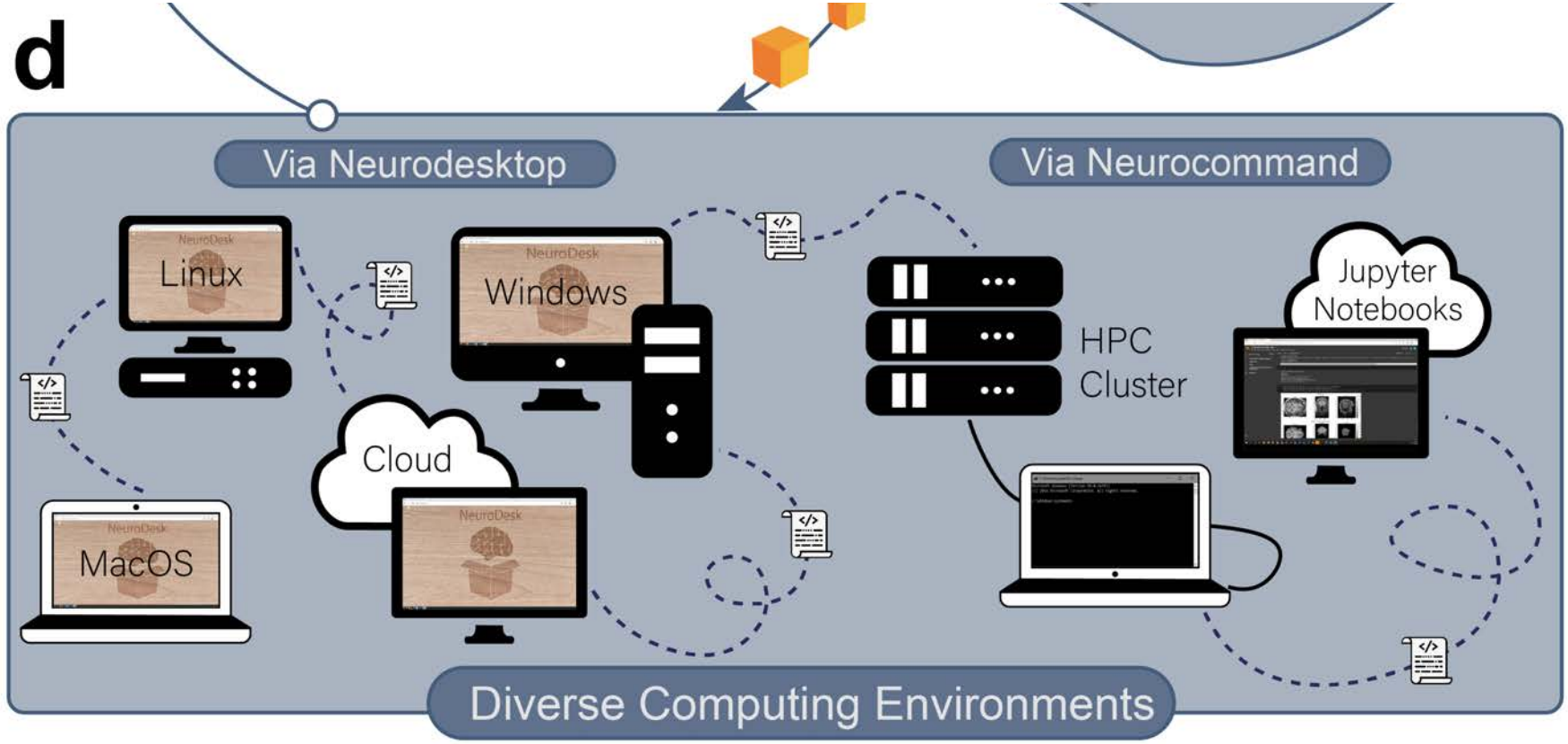
The Neurodesk.org Open-Source Platform

Neurocontainers automatically builds and distributes a repository of software containers



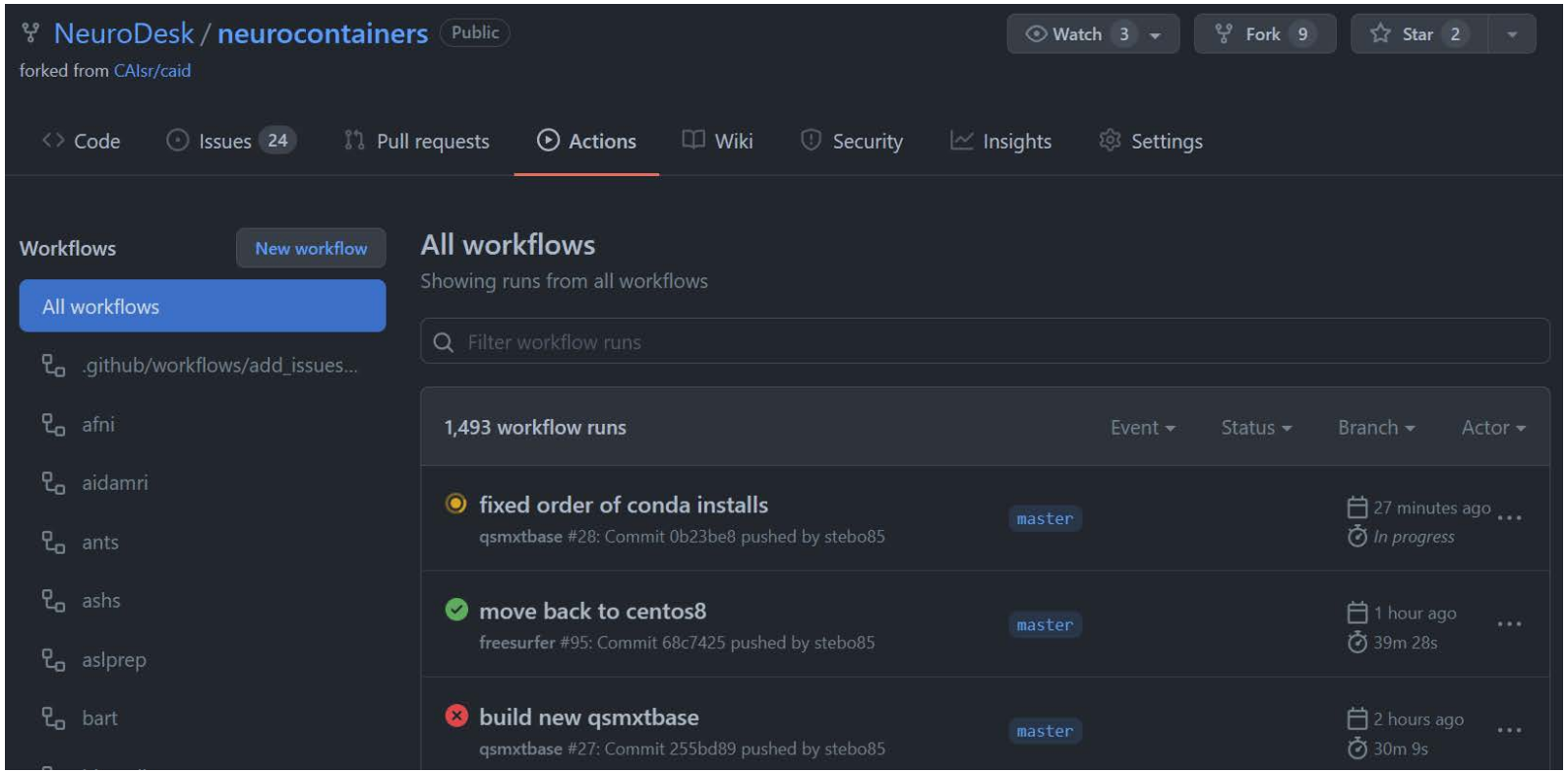
The Neurodesk.org Open-Source Platform

Software containers are available for all compute environments:

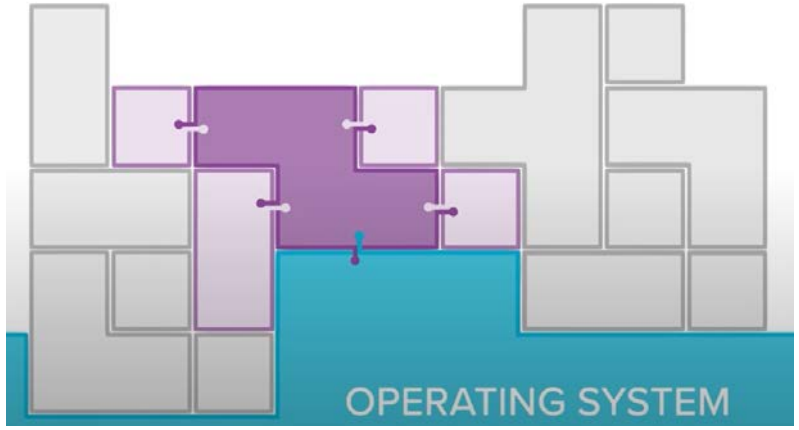


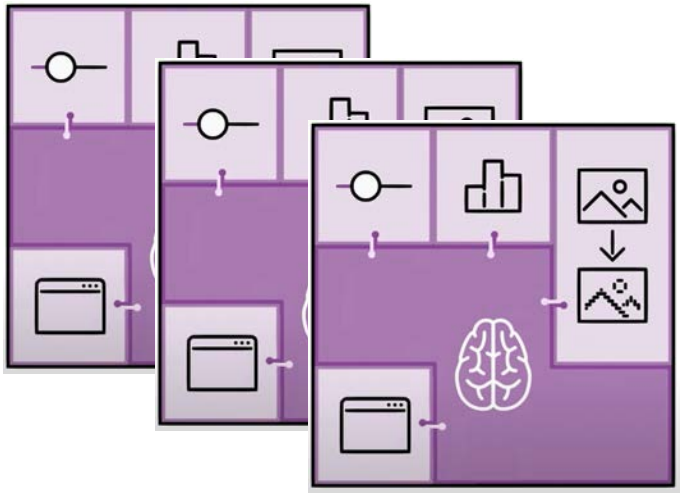
Github actions build and upload the application containers -- all automated

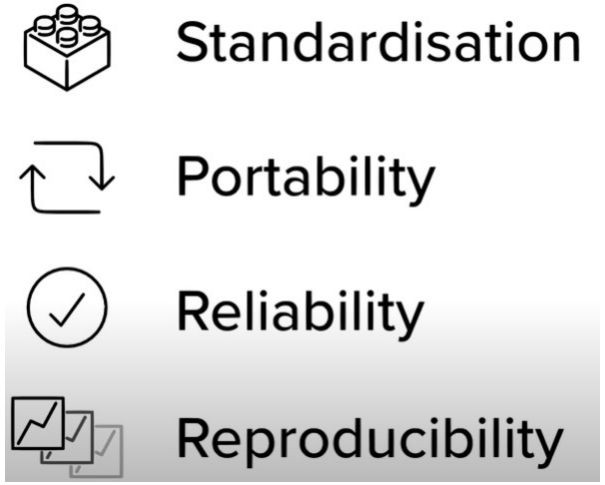
```
#!/usr/bin/env bash
set -e

export toolName='minc'
export toolVersion=1.9.17
# Don't forget to update version change in README.md!!!!!

if [ "$1" != "" ]; then
    echo "Entering Debug mode"
    export debug=$1
fi

source ../main_setup.sh

neurodocker generate ${neurodocker_buildMode} \
   --base-image ubuntu:18.04 \
   --pkg-manager apt \
   --run="mkdir ${mountPointList}" \
   --copy README.md /README.md \
   --${toolName} version=${toolVersion} \
   --env DEPLOY_PATH=/opt/${toolName}-${toolVersion}/bin/:/opt
  > ${imageName}.${neurodocker_buildExt}


if [ "$1" != "" ]; then
   ./../main_build.sh
fi
```

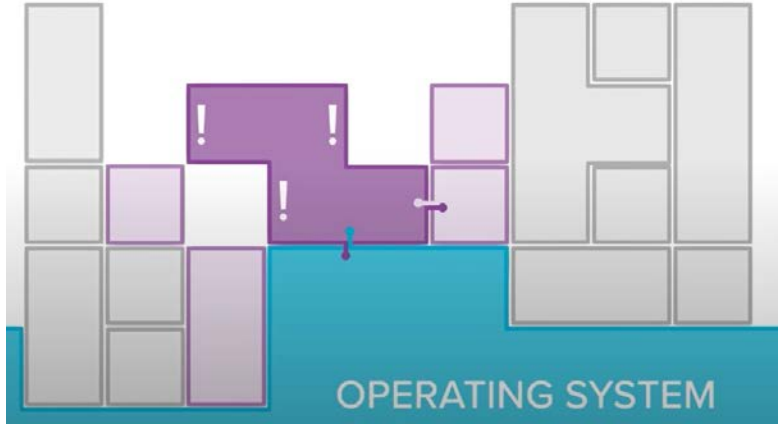


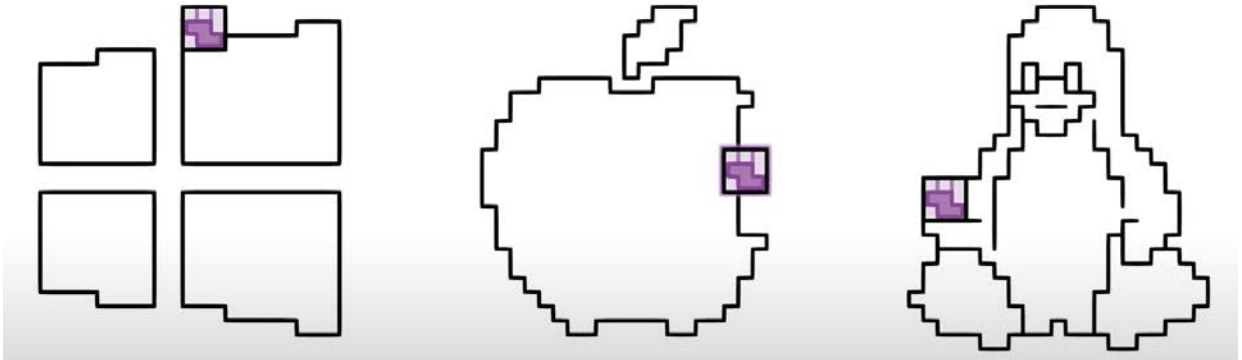
What is a software container, and how can it help?





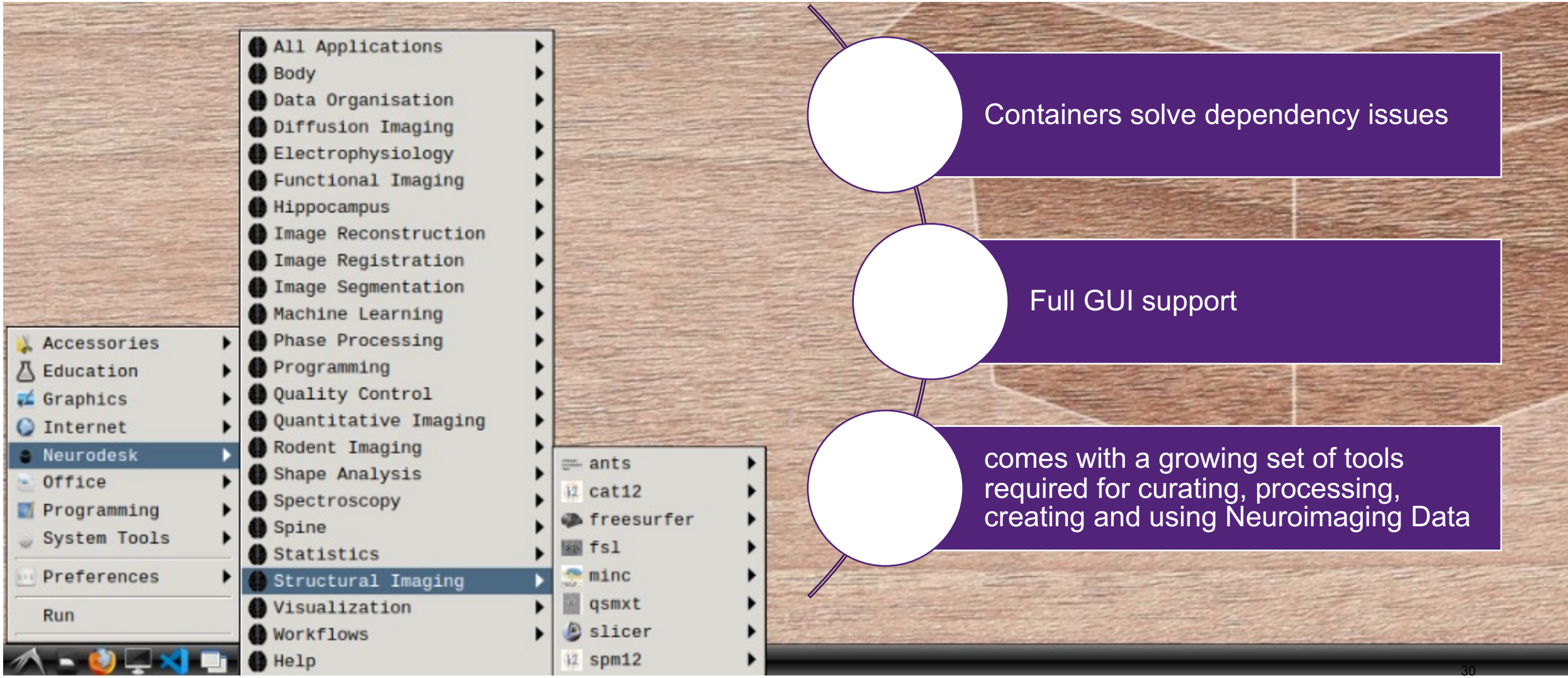






https://www.youtube.com/watch?v=HelrQnm3v4g&ab_channel=AustralianResearchDataCommons-ARDC

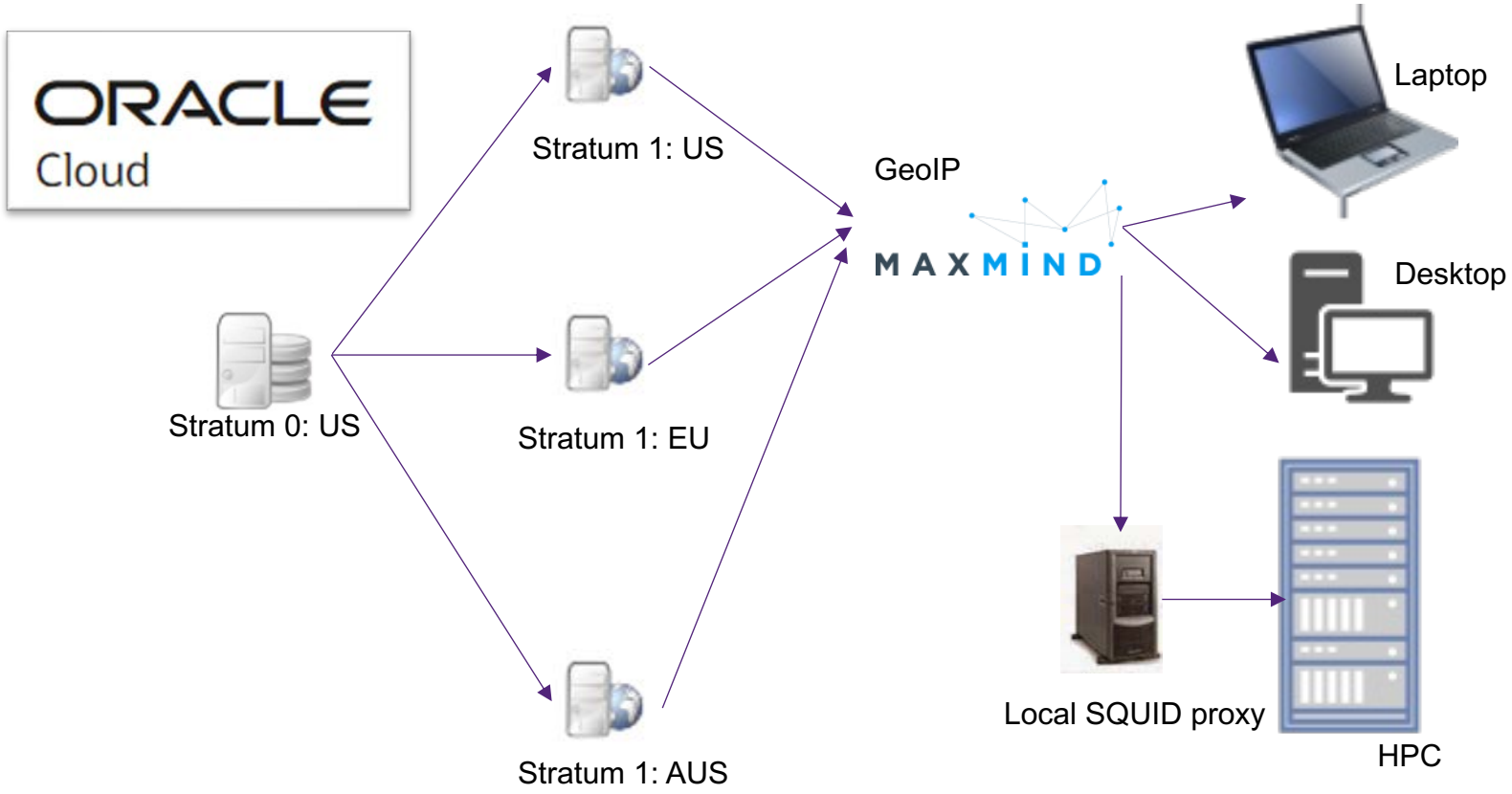
Neurodesktop – A Linux desktop accessible via the browser



- Containers solve dependency issues
- Full GUI support
- comes with a growing set of tools required for curating, processing, creating and using Neuroimaging Data

High performance software distribution using CVMFS

- download and unpack singularity/apptainer containers to CVMFS storage for distribution and on-demand access

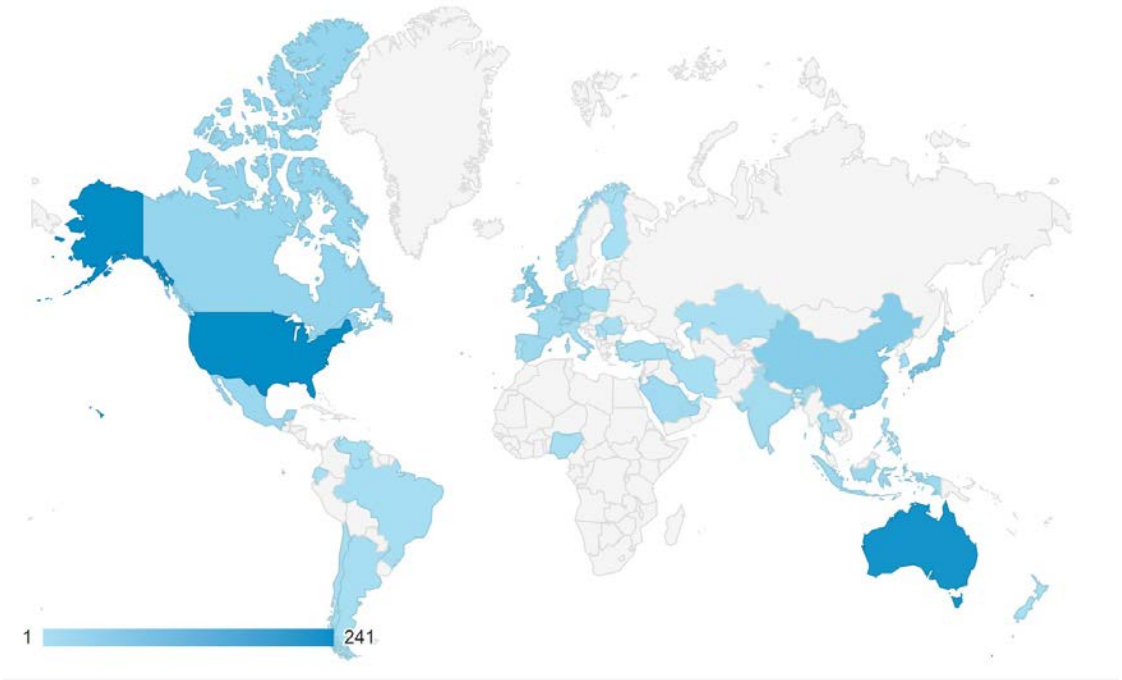



```
ll /cvmfs/neurodesk.ardc.edu.au/containers/

Aug 26 16:41 afni_21.2.00_20210714
Sep 29 12:35 aidamri_1.1_20210708
Aug 26 16:44 aidmri_1.1_20210330
Aug 26 16:49 ants_2.3.4_20210104
Aug 26 17:04 ashs_2.0.0_20210105
Aug 26 16:54 aslprep_0.2.7_20210323
Aug 26 17:05 bart_0.7.00_20210302
Aug 26 17:06 bidstools_1.0.0_20201208
Oct 18 18:01 clearswi_1.0.0_20211018
Oct 13 09:37 condaenvs_1.0.0_20211011
```

https://neurodesk.org/developers/cvmfs/

Uptake in the community

- **996 individual users** in the last 6 months from **47 countries**
- our repos on GitHub have **71 stars & 75 forks**
- development driven by hackathons (e.g. Brainhack Global)



Reproducible Science

Are Neurodesk results reproducible?

How close are we to re-executable papers?

Let's try to reproduce …

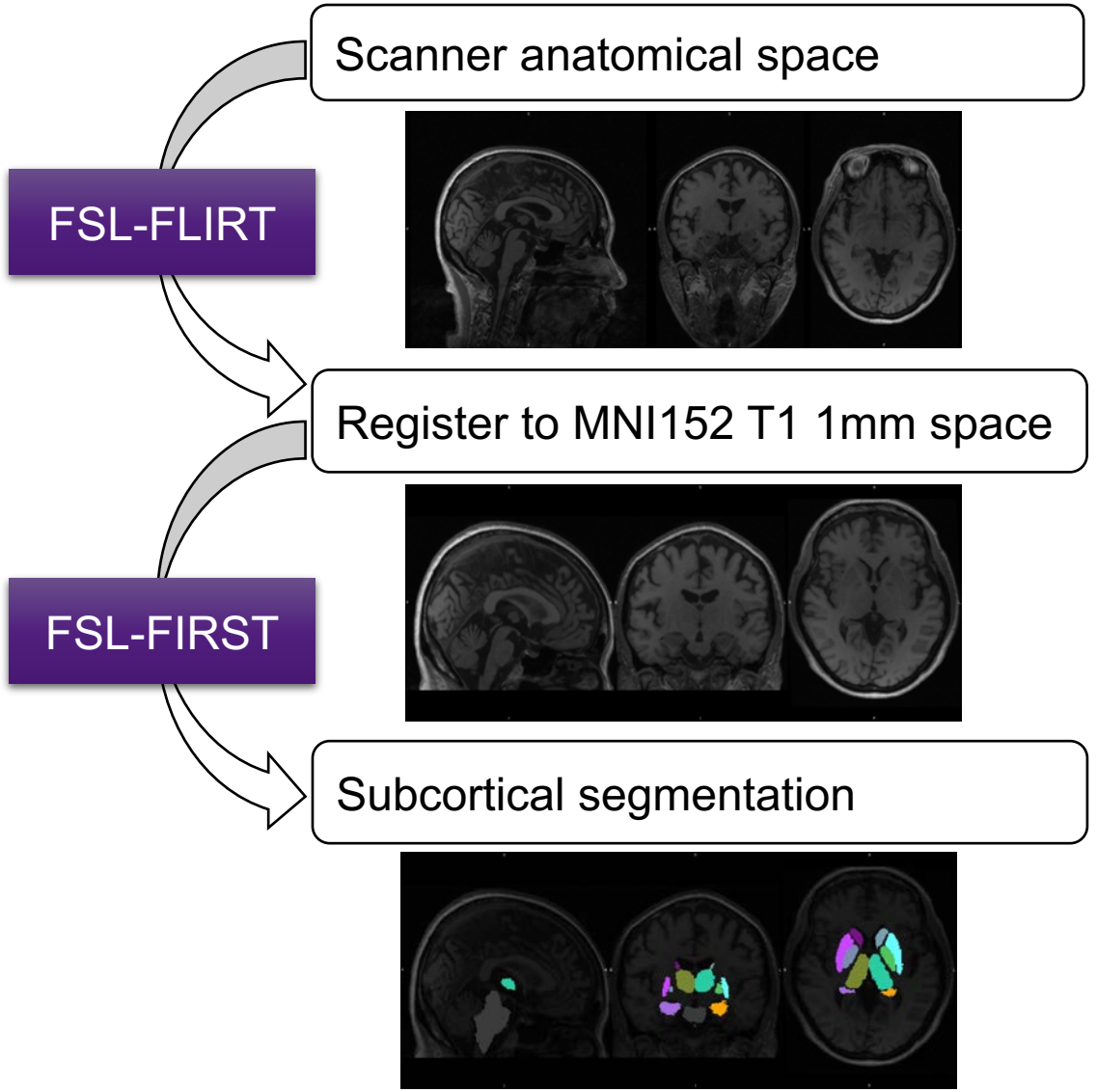
Reproducibility of neuroimaging analyses across operating systems

Tristan Glatard[1,2], Lindsay B. Lewis[1], Rafael Ferreira da Silva[3], Reza Adalat[1], Natacha Beck[1], Claude Lepage[1], Pierre Rioux[1], Marc-Etienne Rousseau[1], Tarek Sherif[1], Ewa Deelman[3], Najmeh Khalili-Mahani[1] and Alan C. Evans[1]

The following analyses were performed by **Thanh Thuy Dao** based on Glatard et al. approach using Neurodesk ☺

The Analysis Setup

	System A		System B	
	Local	Neurodesk	Local	Neurodesk
Applications	FSL 6.0.5.1	FSL 6.0.5.1	FSL 6.0.5.1	FSL 6.0.5.1
Glibc version	2.31	2.23	2.28	2.23
OS	Ubuntu 20.04	Ubuntu 16.04.7	AlmaLinux 8.5	Ubuntu 16.04.7
Hardware	12th Gen Intel(R) Core(TM) i7-12700		AMD EPYC 7542 32-Core Processor	



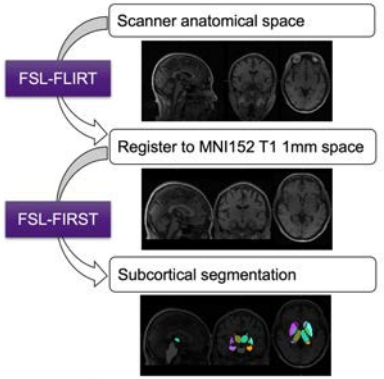
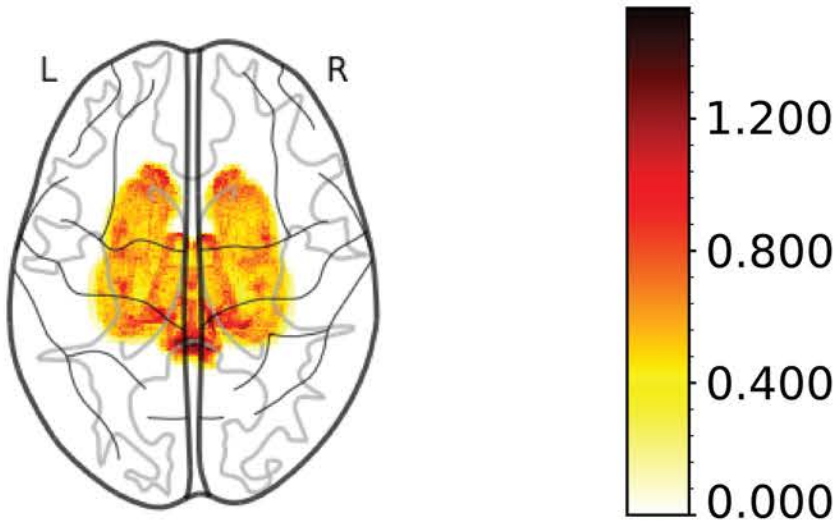


Image Intensity Differences (FSL FLIRT)

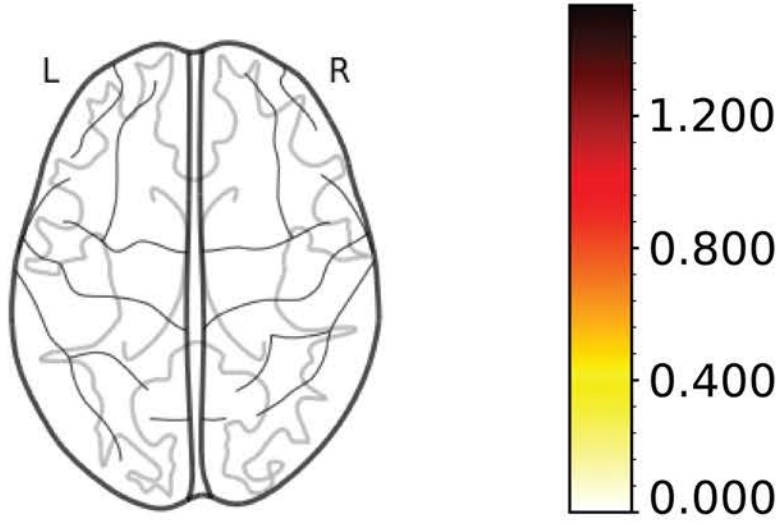
a Local Install

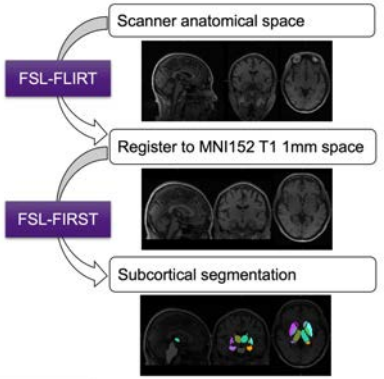
Absolute mean inter-system intensity difference



b Neurodesk

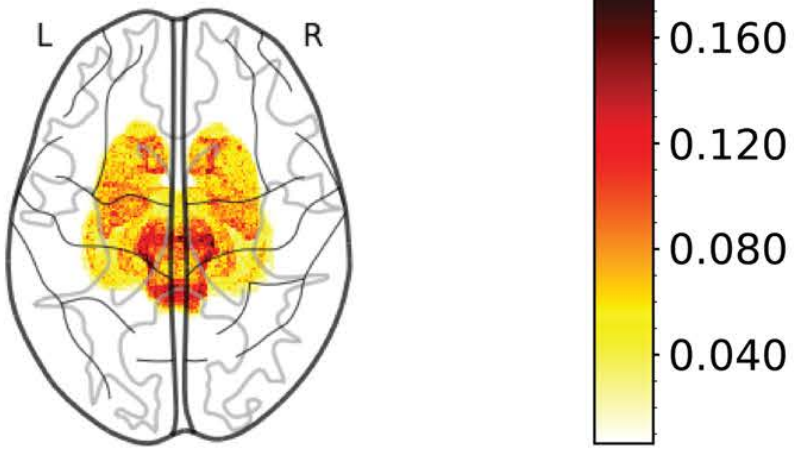
Absolute mean inter-system intensity difference



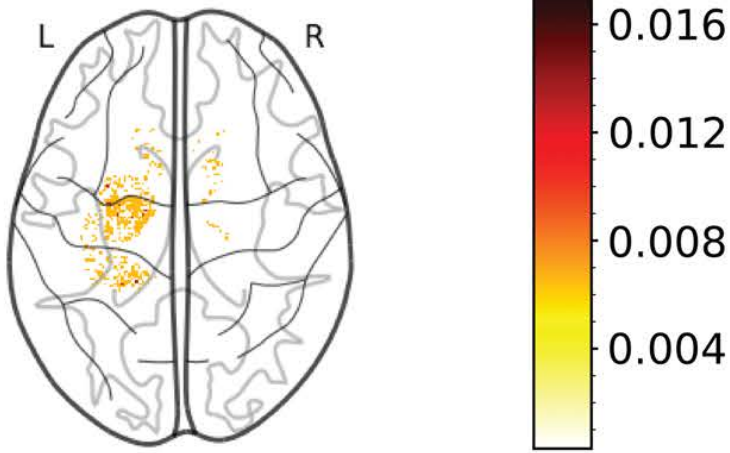


Classification Differences (FSL FIRST)

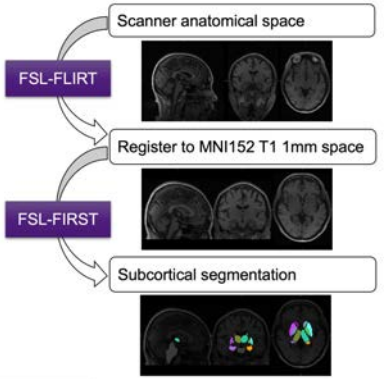
c Local Install

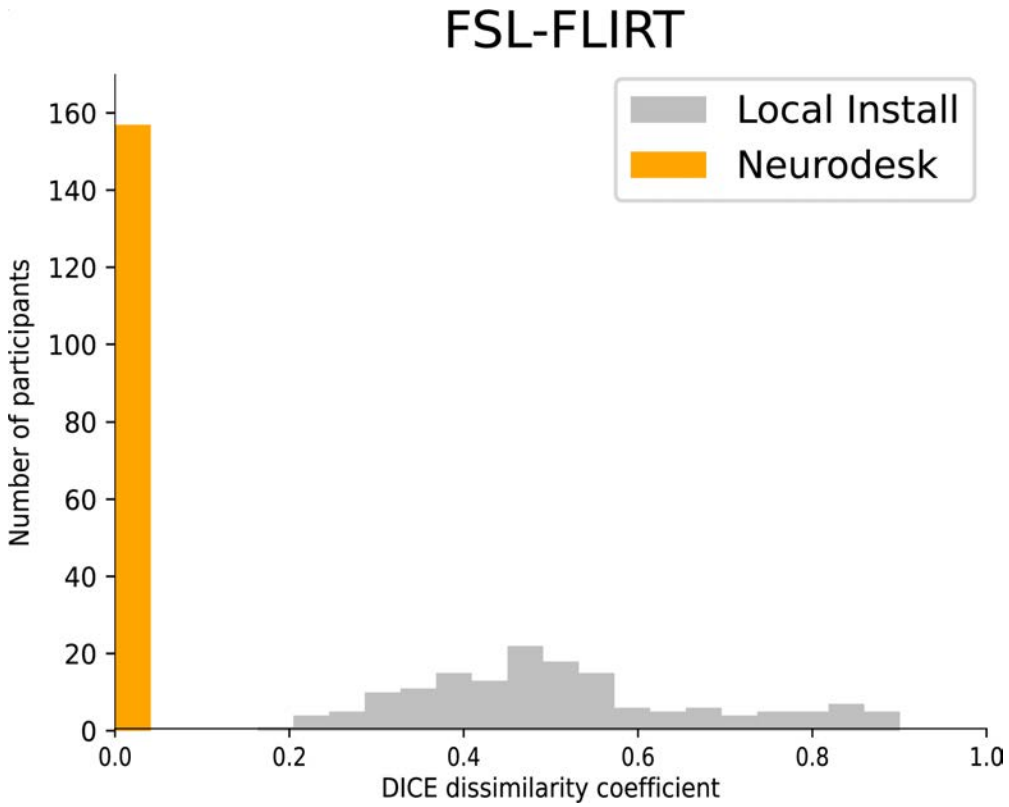


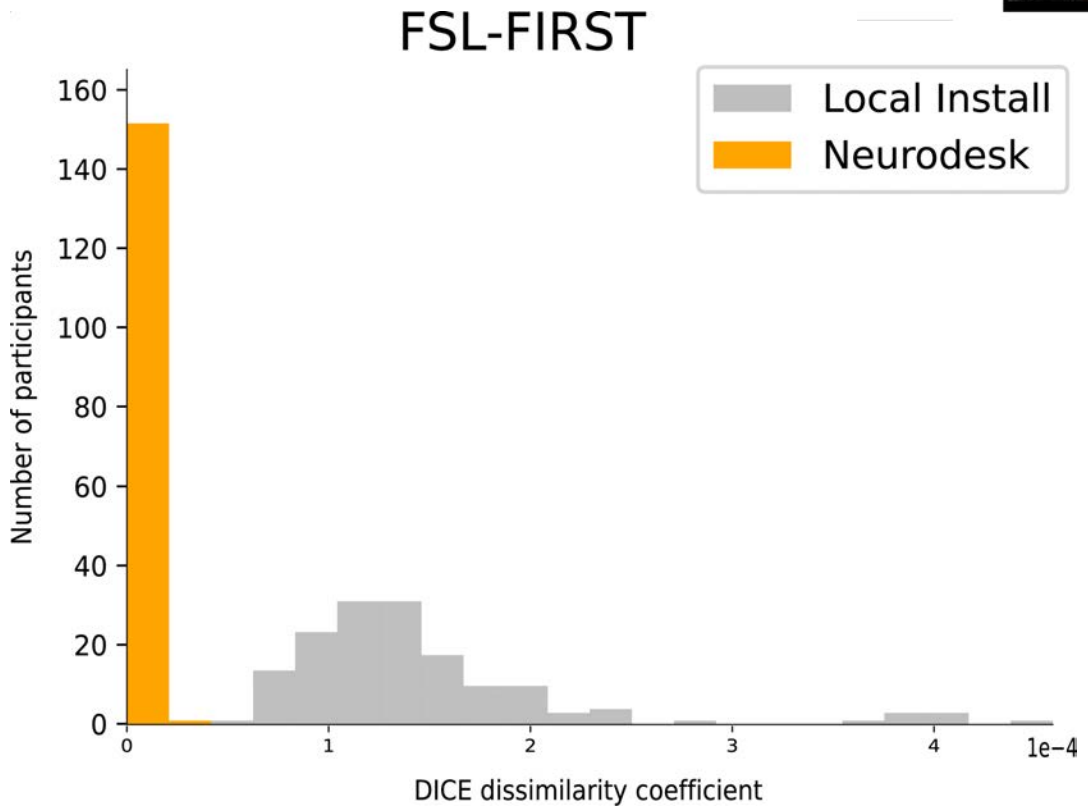
d Neurodesk

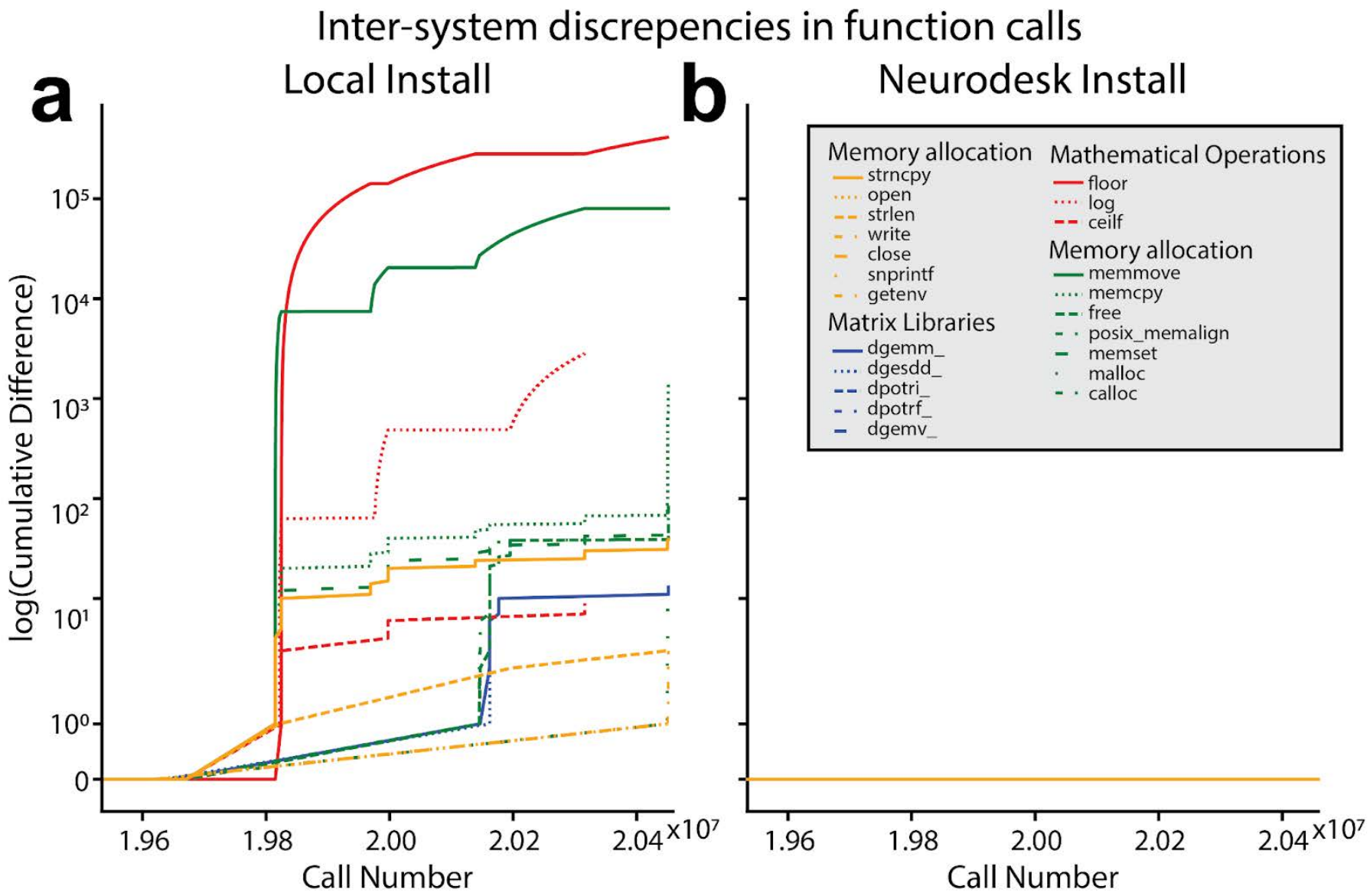


DICE dissimilarity between System A and B








Inter-system discrepencies in function calls
a
Local Install
b
Neurodesk Install
log(Cumulative Difference)
10^5
10^4
10^3
10^2
10^1
10^0
0
1.96
1.98
2.00
2.02
2.04
x10⁷
Call Number
Memory allocation
strncpy
open
strlen
write
close
snprintf
getenv
Matrix Libraries
dgemm_
dgesdd_
dpotri_
dpotrf_
dgemv_
Mathematical Operations
floor
log
ceilf
Memory allocation
memmove
memcpy
free
posix_memalign
memset
malloc
calloc

Interactive papers

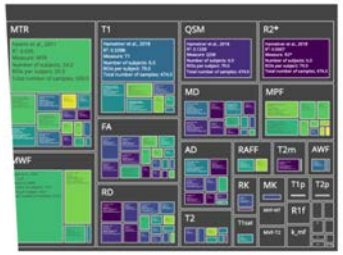
CONP PCNO

Supported by the Canadian Open Neuroscience Platform (CONP).

The quest for measuring myelin with MRI - An interactive meta-analysis

This study explores an important aspect of quantitative magnetic resonance imaging (qMRI): validation. Focusing specifically on myelin measures, we show the results of our meta-analysis comparing quantitative MRI with histology.

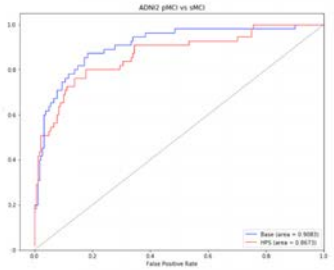
NeuroLibre Book GitHub Code



A highly predictive signature (HPS) of Alzheimer's disease dementia from cognitive and structural brain features

A jupyter notebook containing analyses that give a highly predictive signature (HPS) of Alzheimer's disease dementia from cognitive and structural features using simulated data.

NeuroLibre Book GitHub Code



https://www.neurolibre.com/

eLife

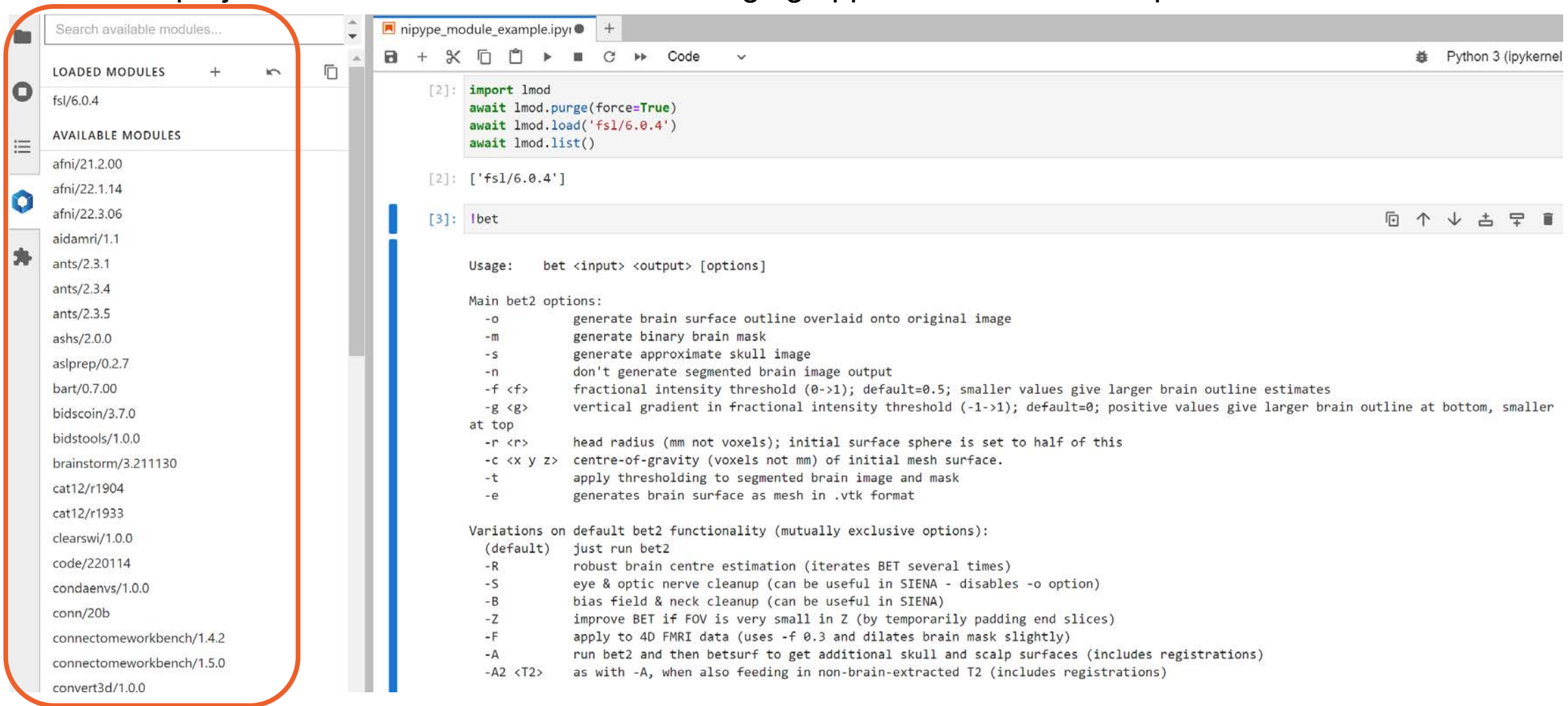
Steps to enrich your published article with code and data

1. Convert your eLife article into R Markdown/ Jupyter notebook with Stencila Hub
2. Add code chunks into the R Markdown/ Jupyter notebook locally
3. Upload enriched R Markdown/ Jupyter notebook and necessary tabular data files to Stencila Hub
4. Share a project snapshot with the eLife Production Team

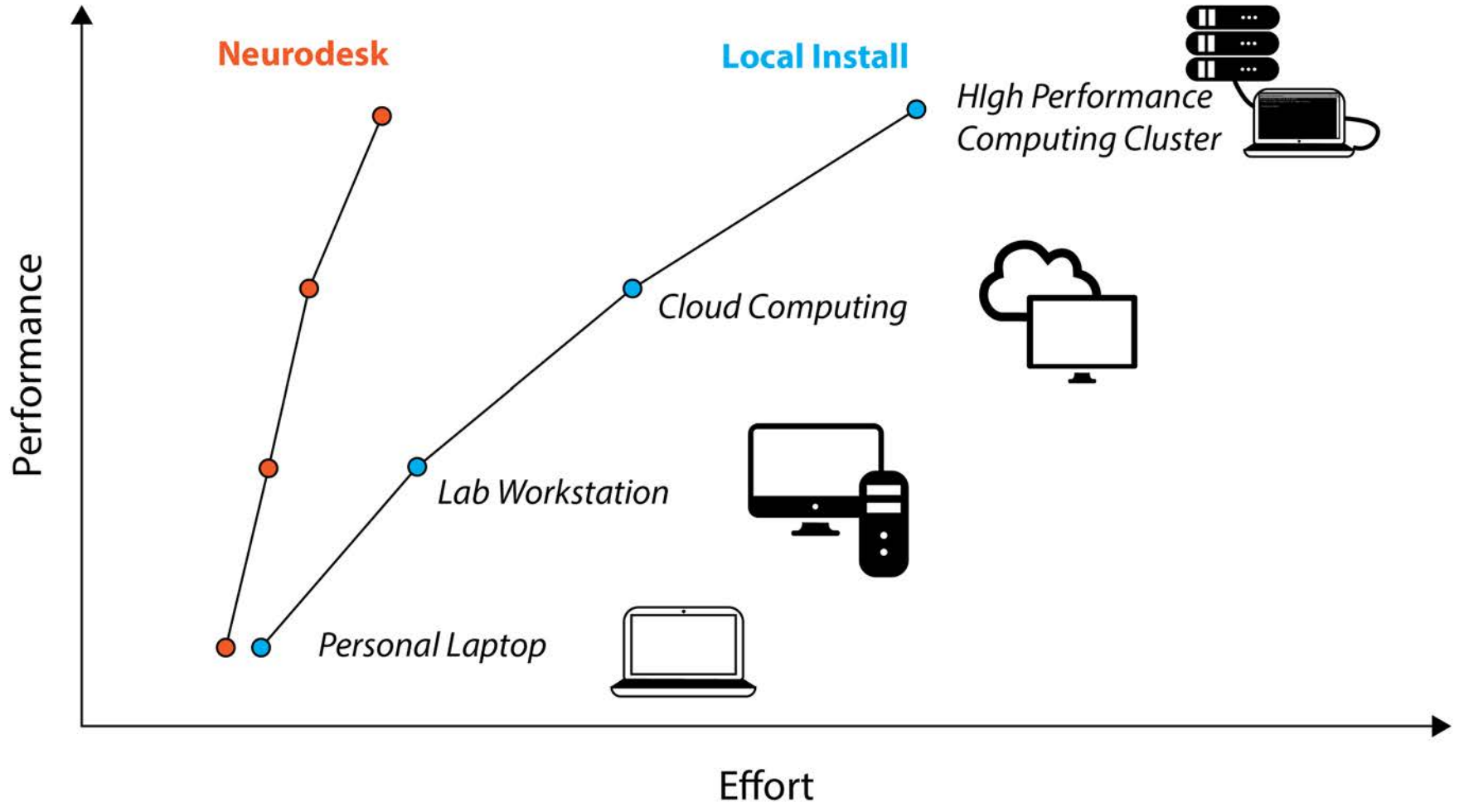
https://elifesciences.org/labs/dc5acbde/welcome-to-a-new-era-of-reproducible-publishing

Neurodesk applications & Jupyter notebooks

Neurodesk project enables the use of all neuroimaging applications inside computational notebooks:



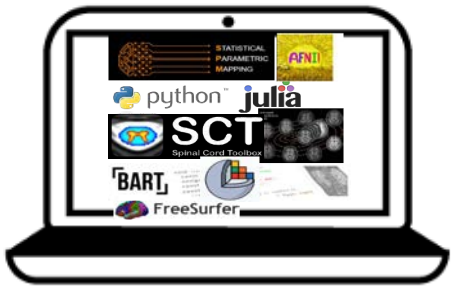
Neurodesk at Scale



Neurodesk brings …

… a suite of neuroimaging tools …

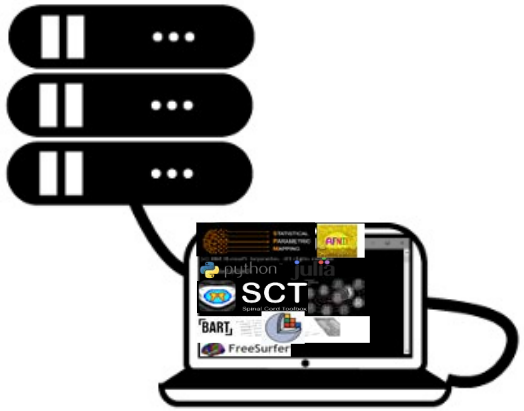
… on your notebook!



… on your lab workstation!

… on a cloud provider!

… on the university's high performance cluster!



Roadmap

More Neuroimaging containers

Support of GPUs

Support of M1/Arm processors

Federated Learning For Segmentation

MONAI+

A fully javascript/ webassembly based version -> no docker dependencies

WA