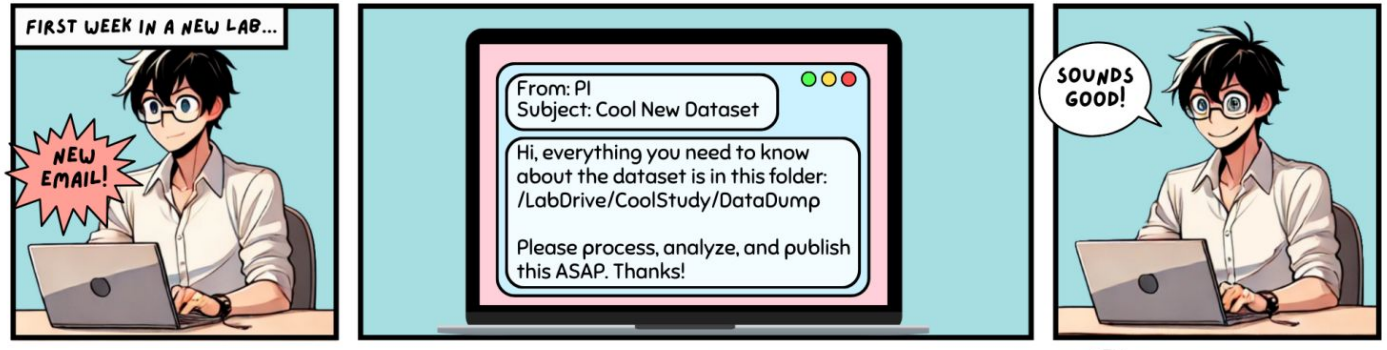
Assessing analytical variability with **Nipoppy**

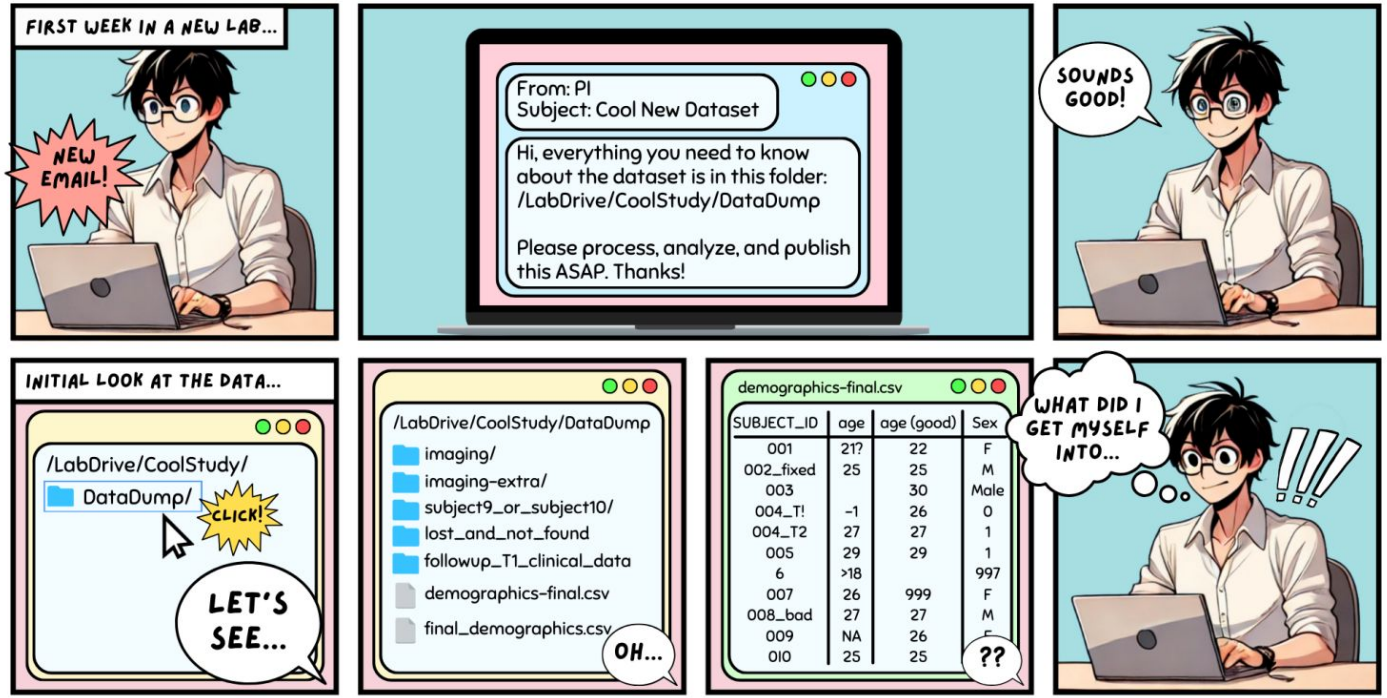
Michelle Wang
NeuroDataScience-ORIGAMI Lab (JB Poline)
McGill University

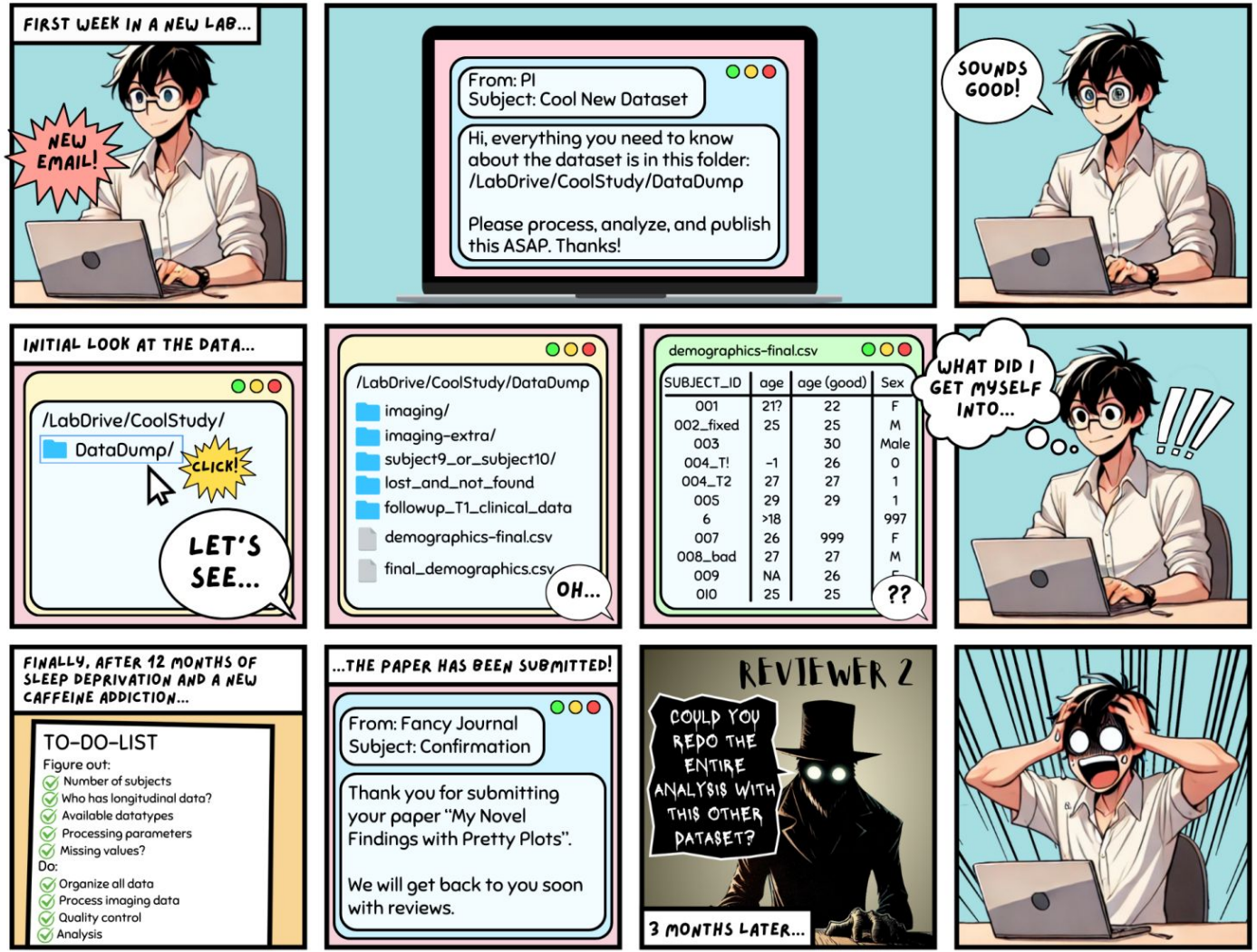
2025 June 24

ORIGAMI
Lab

OHBM 2025


FIRST WEEK IN A NEW LAB...
NEW EMAIL!
From: PI
Subject: Cool New Dataset
Hi, everything you need to know about the dataset is in this folder: /LabDrive/CoolStudy/DataDump
Please process, analyze, and publish this ASAP. Thanks!
SOUNDS GOOD!


FIRST WEEK IN A NEW LAB...
NEW EMAIL!
From: PI
Subject: Cool New Dataset
Hi, everything you need to know about the dataset is in this folder: /LabDrive/CoolStudy/DataDump
Please process, analyze, and publish this ASAP. Thanks!
SOUNDS GOOD!
INITIAL LOOK AT THE DATA...
/LabDrive/CoolStudy/
DataDump/
CLICK!
LET'S SEE...
/LabDrive/CoolStudy/DataDump
imaging/
imaging-extra/
subject9_or_subject10/
lost_and_not_found
followup_T1_clinical_data
demographics-final.csv
final_demographics.csv
OH...
demographics-final.csv
SUBJECT_ID | age | age (good) | Sex
001 | 21? | 22 | F
002_fixed | 25 | 25 | M
003 | | 30 | Male
004_T! | -1 | 26 | 0
004_T2 | 27 | 27 | 1
005 | 29 | 29 | 1
6 | >18 | | 997
007 | 26 | 999 | F
008_bad | 27 | 27 | M
009 | NA | 26 |
010 | 25 | 25 |
??
WHAT DID I GET MYSELF INTO...
!!!


FIRST WEEK IN A NEW LAB...
NEW EMAIL!
From: PI
Subject: Cool New Dataset
Hi, everything you need to know about the dataset is in this folder: /LabDrive/CoolStudy/DataDump
Please process, analyze, and publish this ASAP. Thanks!
SOUNDS GOOD!
INITIAL LOOK AT THE DATA...
/LabDrive/CoolStudy/
DataDump/
CLICK!
LET'S SEE...
/LabDrive/CoolStudy/DataDump
imaging/
imaging-extra/
subject9_or_subject10/
lost_and_not_found
followup_T1_clinical_data
demographics-final.csv
final_demographics.csv
OH...
demographics-final.csv
SUBJECT_ID | age | age (good) | Sex
001 | 21? | 22 | F
002_fixed | 25 | 25 | M
003 | | 30 | Male
004_T! | -1 | 26 | 0
004_T2 | 27 | 27 | 1
005 | 29 | 29 | 1
6 | >18 | | 997
007 | 26 | 999 | F
008_bad | 27 | 27 | M
009 | NA | 26 |
010 | 25 | 25 |
??
WHAT DID I GET MYSELF INTO...
!!!
FINALLY, AFTER 12 MONTHS OF SLEEP DEPRIVATION AND A NEW CAFFEINE ADDICTION...
TO-DO-LIST
Figure out:
Number of subjects
Who has longitudinal data?
Available datatypes
Processing parameters
Missing values?
Do:
Organize all data
Process imaging data
Quality control
Analysis
...THE PAPER HAS BEEN SUBMITTED!
From: Fancy Journal
Subject: Confirmation
Thank you for submitting your paper "My Novel Findings with Pretty Plots".
We will get back to you soon with reviews.
REVIEWER 2
COULD YOU REDO THE ENTIRE ANALYSIS WITH THIS OTHER DATASET?
3 MONTHS LATER...

How can we make neuroimaging datasets more FAIR?

- **FAIR**: findable, accessible, interoperable, reusable (Wilkinson *et al.*, 2016)

Existing open standards/tools developed by the community

(Gorgolewski *et al.*, 2016)

(Kurtzer *et al.*, 2017)

(Glatard *et al.*, 2018)

We leveraged existing open science tools to build a flexible framework for data organization and processing of neuroimaging-clinical data

Introducing the **Nipoppy** framework

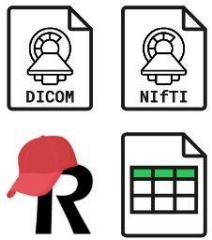
1. **Protocol** for best practices (FAIR principles) in organization/processing of neuroimaging-clinical datasets

Introducing the **Nipoppy** framework

1. **Protocol** for best practices (FAIR principles) in organization/processing of neuroimaging-clinical datasets

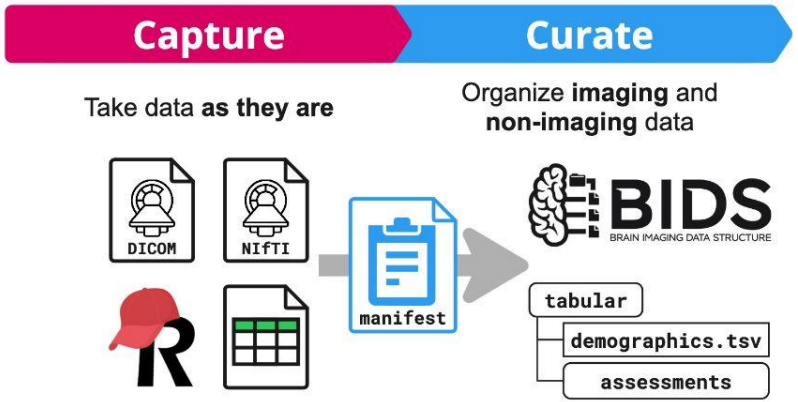
Capture

Take data **as they are**



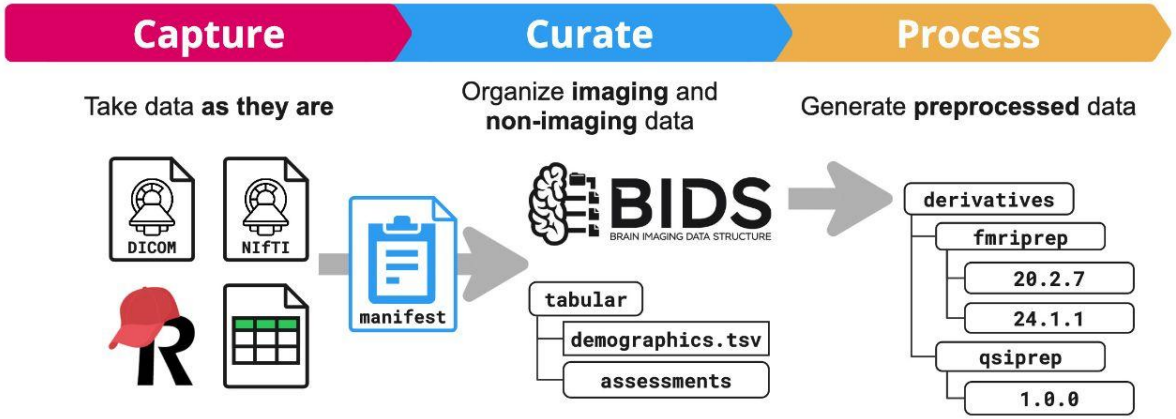
Introducing the **Nipoppy** framework

1. **Protocol** for best practices (FAIR principles) in organization/processing of neuroimaging-clinical datasets



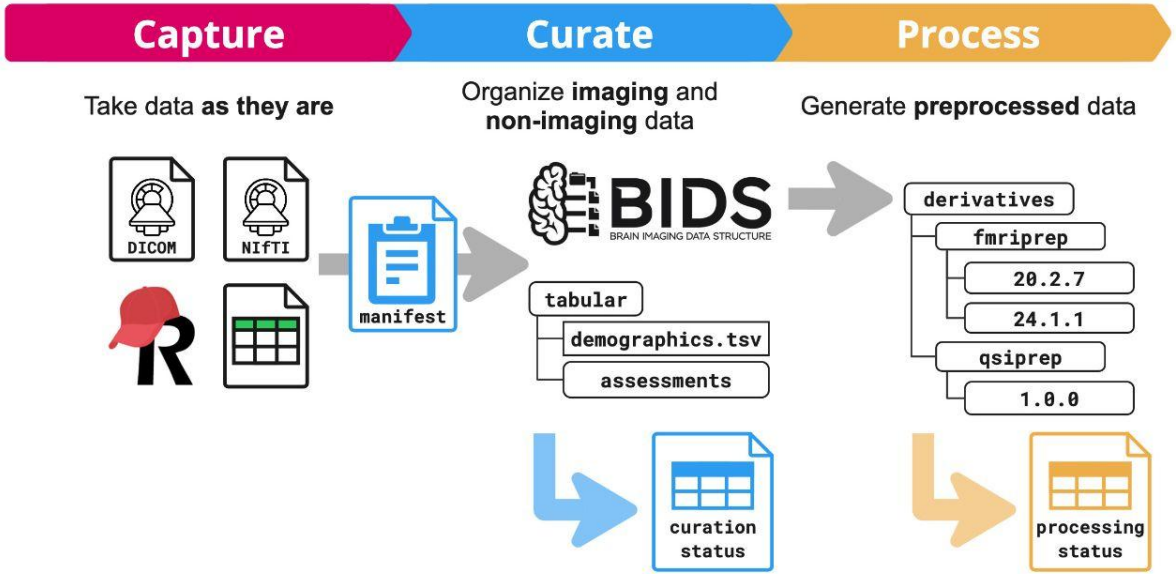
Introducing the **Nipoppy** framework

1. **Protocol** for best practices (FAIR principles) in organization/processing of neuroimaging-clinical datasets



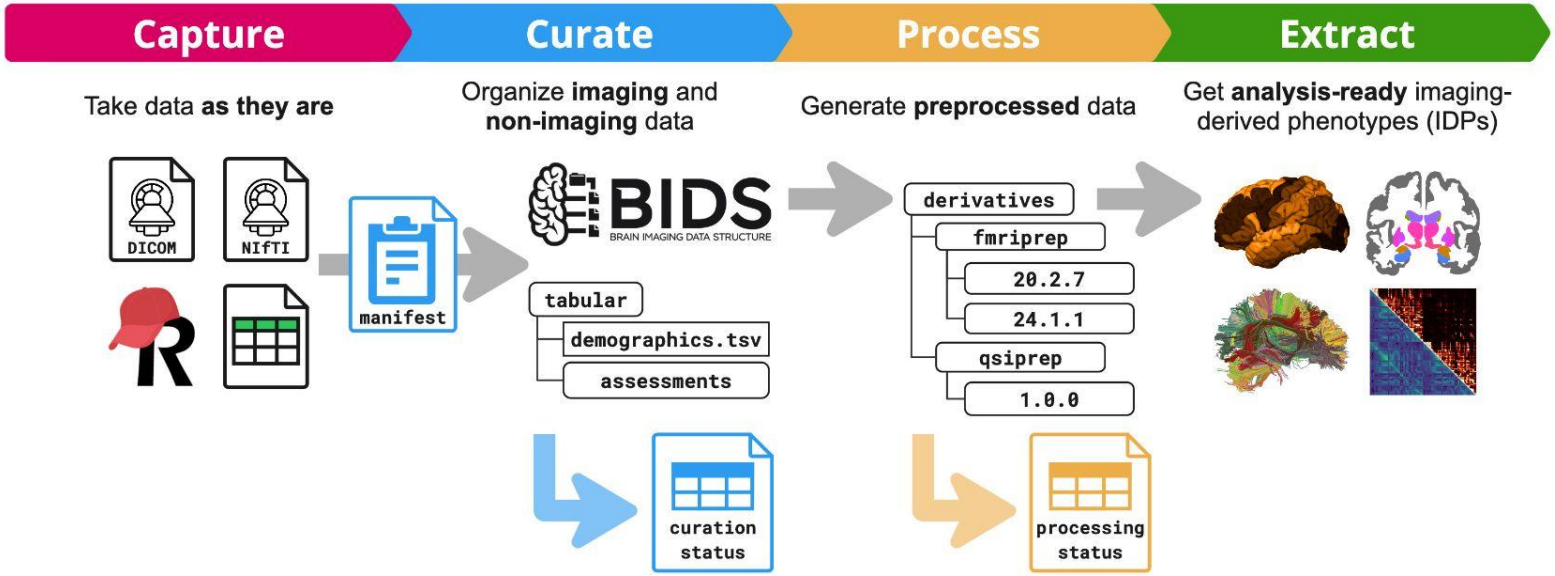
Introducing the **Nipoppy** framework

1. **Protocol** for best practices (FAIR principles) in organization/processing of neuroimaging-clinical datasets



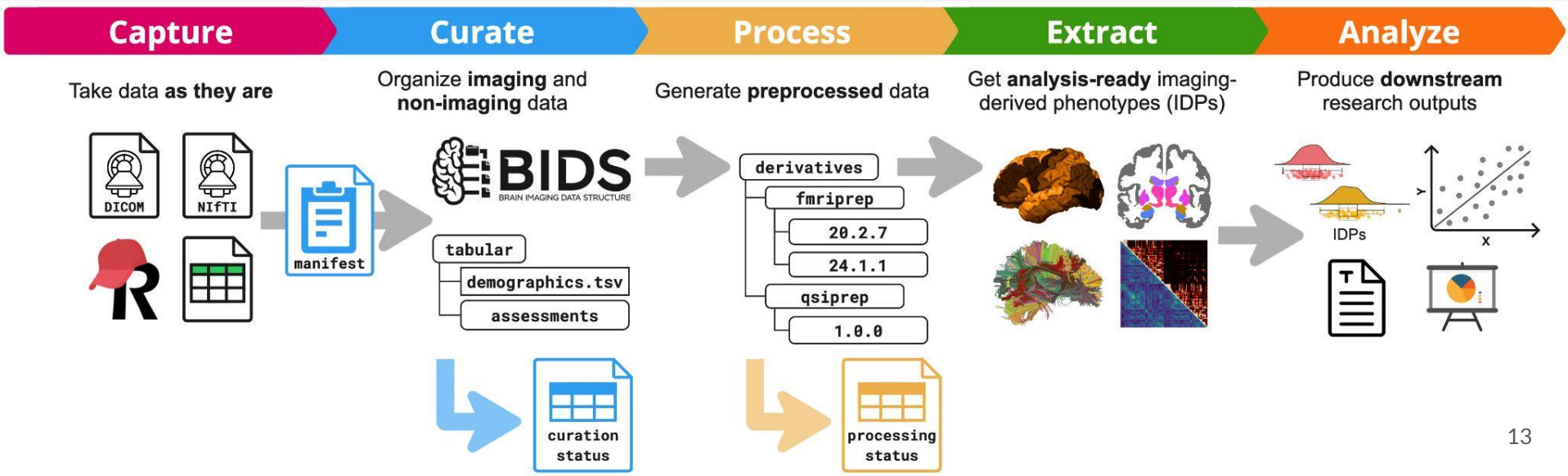
Introducing the **Nipoppy** framework

1. **Protocol** for best practices (FAIR principles) in organization/processing of neuroimaging-clinical datasets



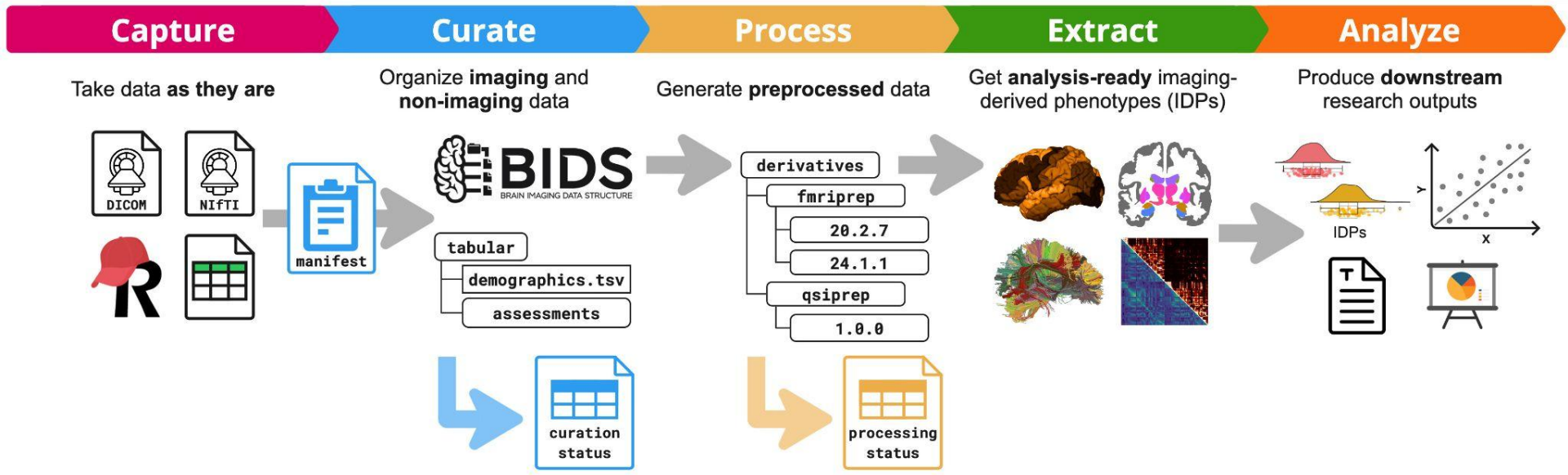
Introducing the **Nipoppy** framework

1. **Protocol** for best practices (FAIR principles) in organization/processing of neuroimaging-clinical datasets

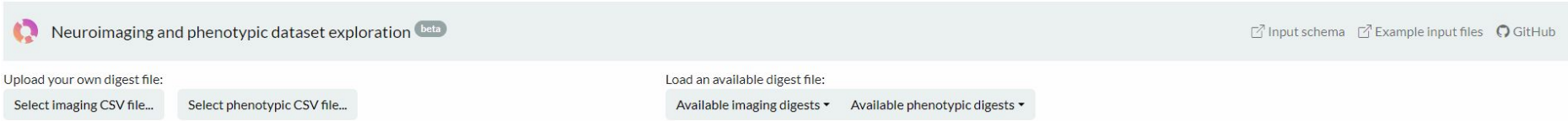


Introducing the **Nipoppy** framework

1. **Protocol** for best practices (FAIR principles) in organization/processing of neuroimaging-clinical datasets
   a. From raw scanner data to extracted imaging-derived phenotypes (IDPs)
   b. Visualizing tabular data/metadata on a dashboard (https://digest.neurobagel.org/)



https://digest.neurobagel.org/



Introducing the **Nipoppy** framework

1. **Protocol** for best practices (FAIR principles) in organization/processing of neuroimaging-clinical datasets
   a. From raw scanner data to extracted imaging-derived phenotypes (IDPs)
   b. Visualizing tabular data/metadata on a dashboard
      (https://digest.neurobagel.org/)

Introducing the **Nipoppy** framework

1. **Protocol** for best practices (FAIR principles) in organization/processing of neuroimaging-clinical datasets
   a. From raw scanner data to extracted imaging-derived phenotypes (IDPs)
   b. Visualizing tabular data/metadata on a dashboard
   (https://digest.neurobagel.org/)

2. Data organization **specification**
   a. For imaging and non-imaging data
   b. At the whole study level

Introducing the **Nipoppy** framework

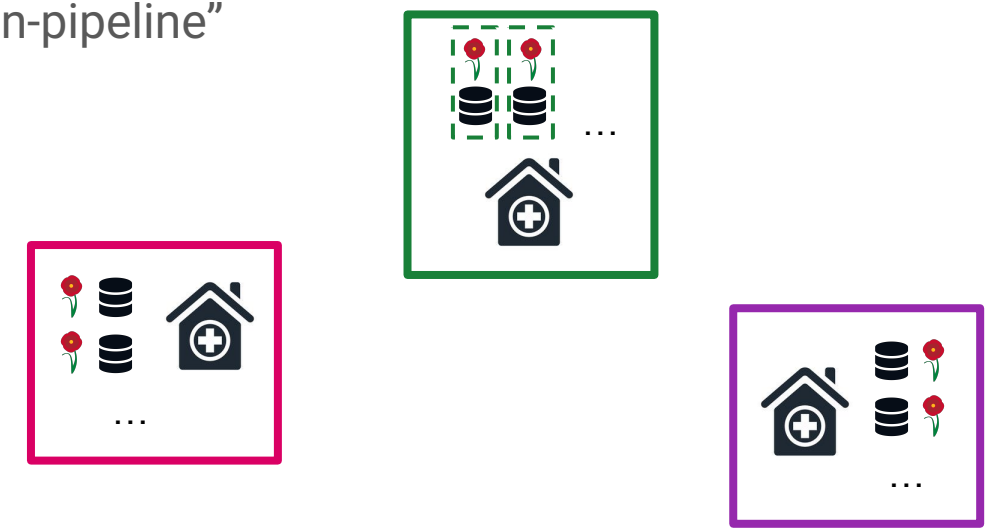
1. **Protocol** for best practices (FAIR principles) in organization/processing of neuroimaging-clinical datasets
   a. From raw scanner data to extracted imaging-derived phenotypes (IDPs)
   b. Visualizing tabular data/metadata on a dashboard
   (https://digest.neurobagel.org/)

2. Data organization **specification**
   a. For imaging and non-imaging data
   b. At the whole study level

3. Software package with **tools** to help work with this framework **locally**

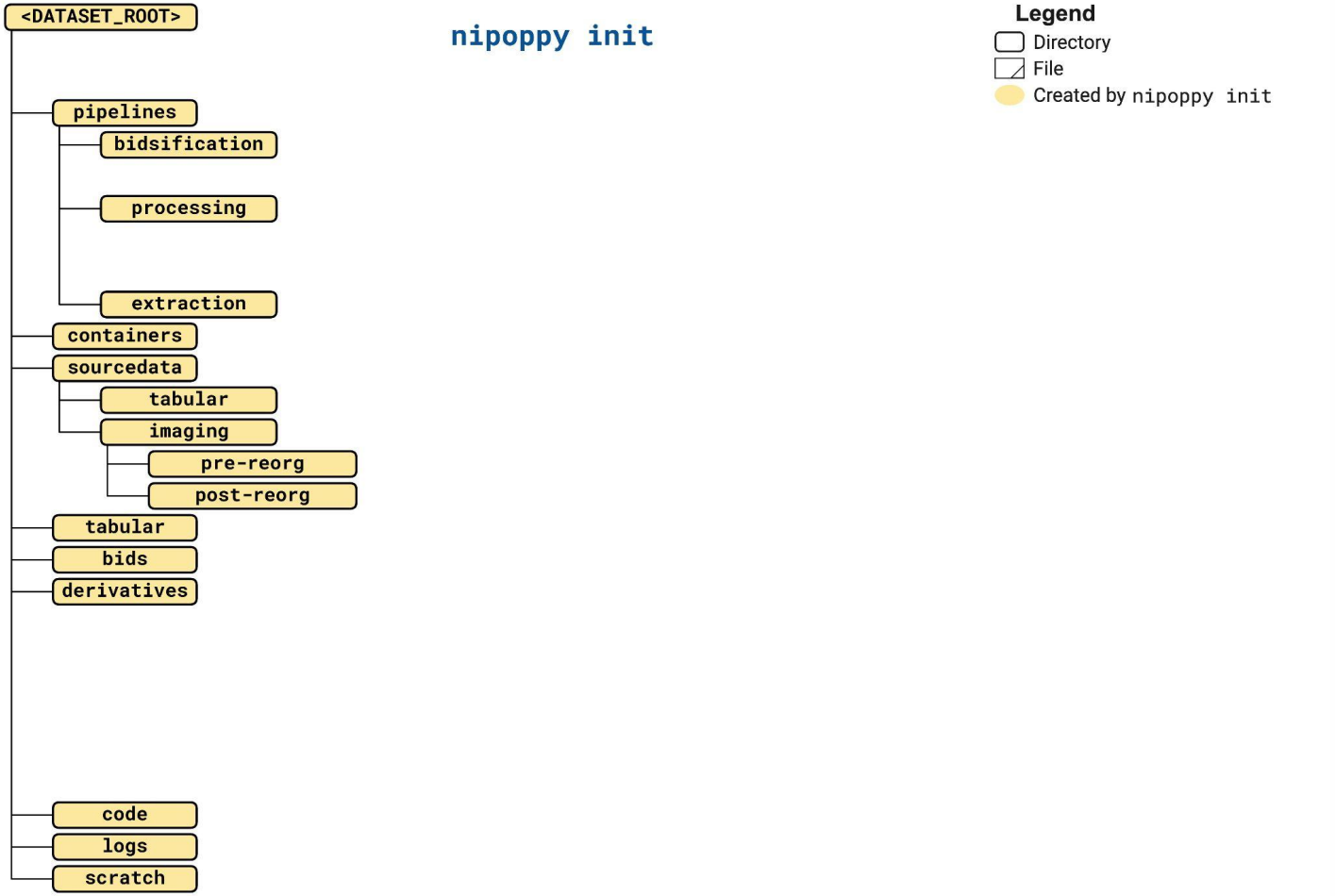
Nipoppy for **decentralized** data management/processing

- `pip`-installable Python package
- Local processing
- Good for consortium-type settings
  - **Independent but consistent** processing (same pipelines/parameters)
- “Bring-your-own-pipeline”

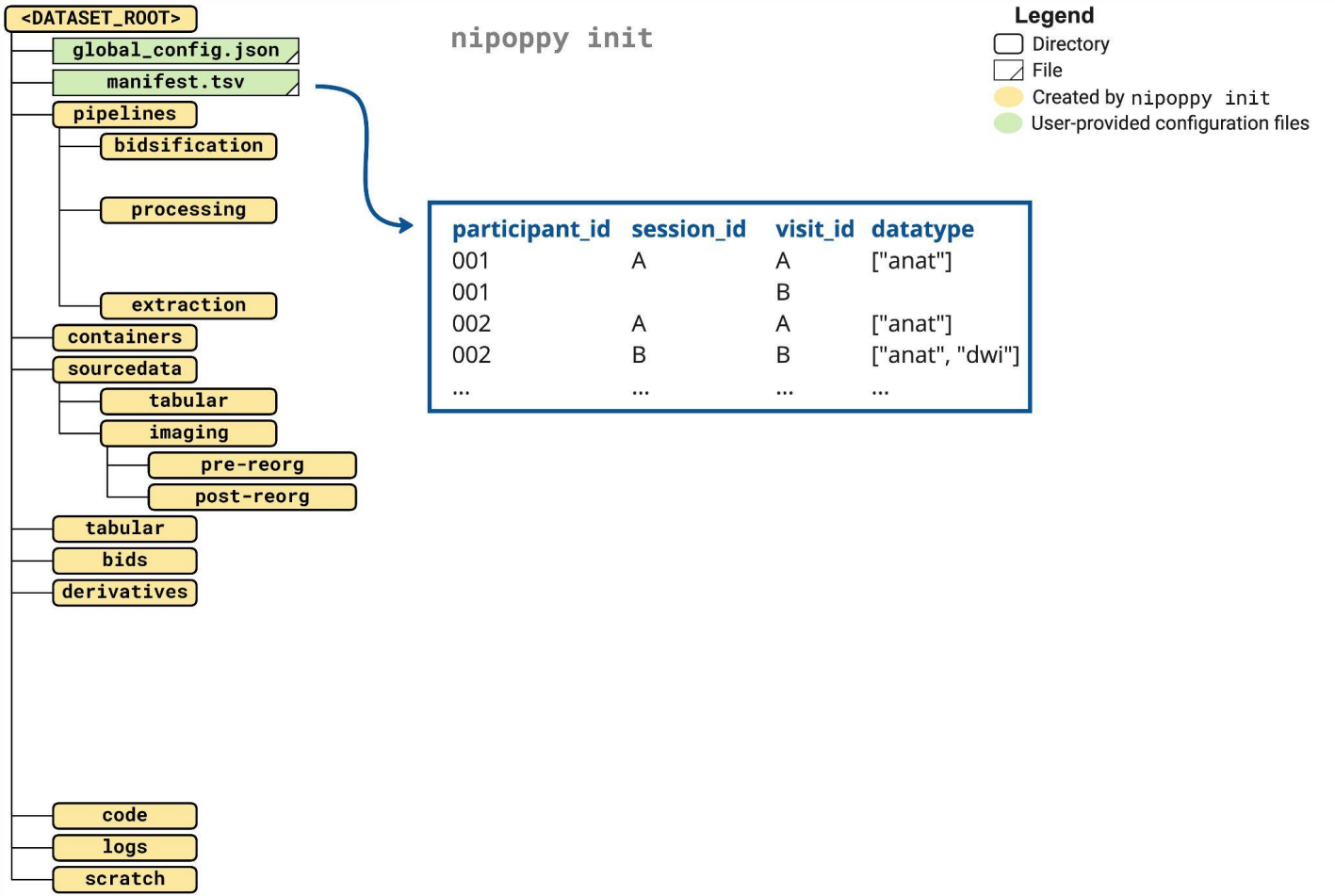


A typical **Nipoppy** workflow

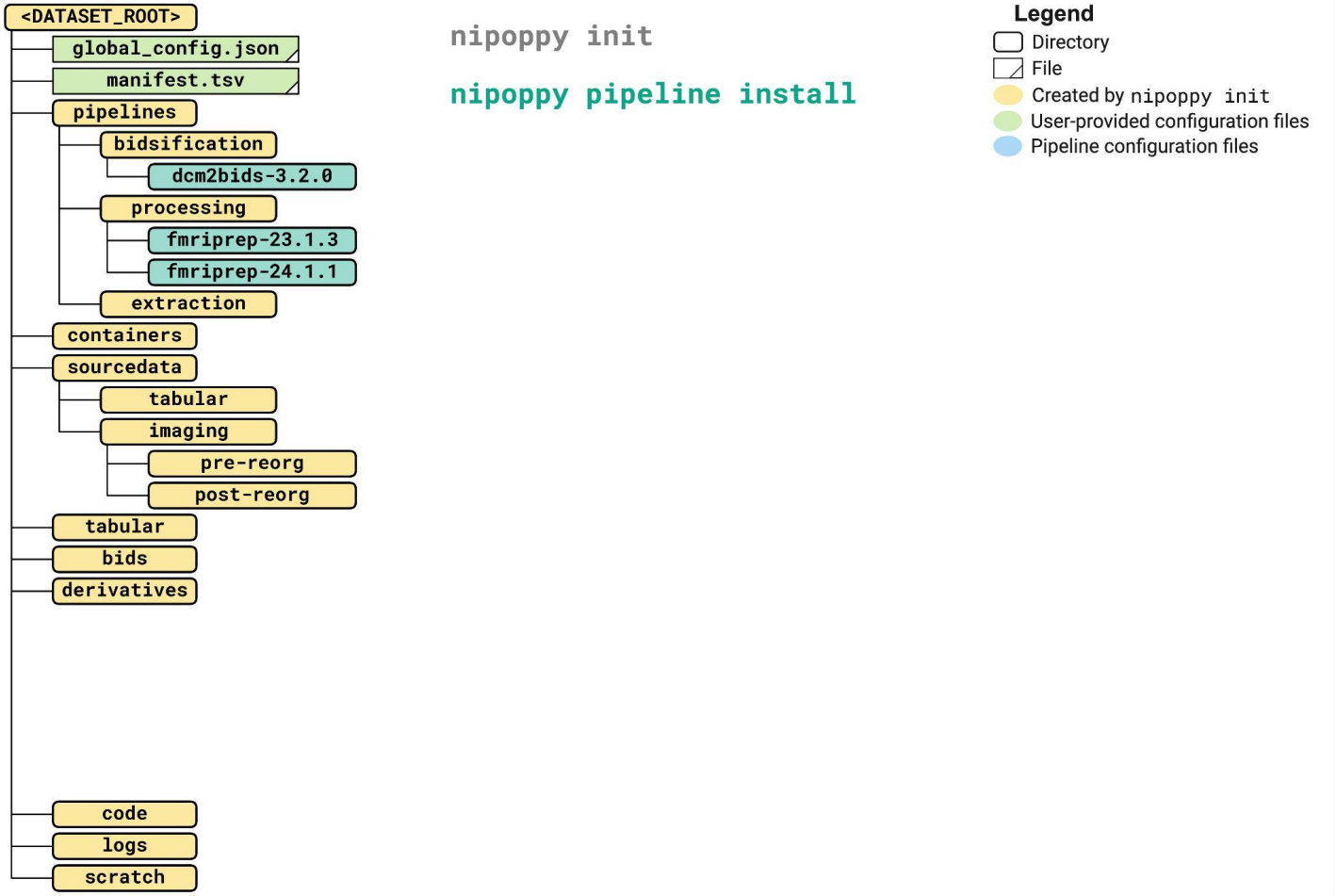
A typical Nipoppy workflow



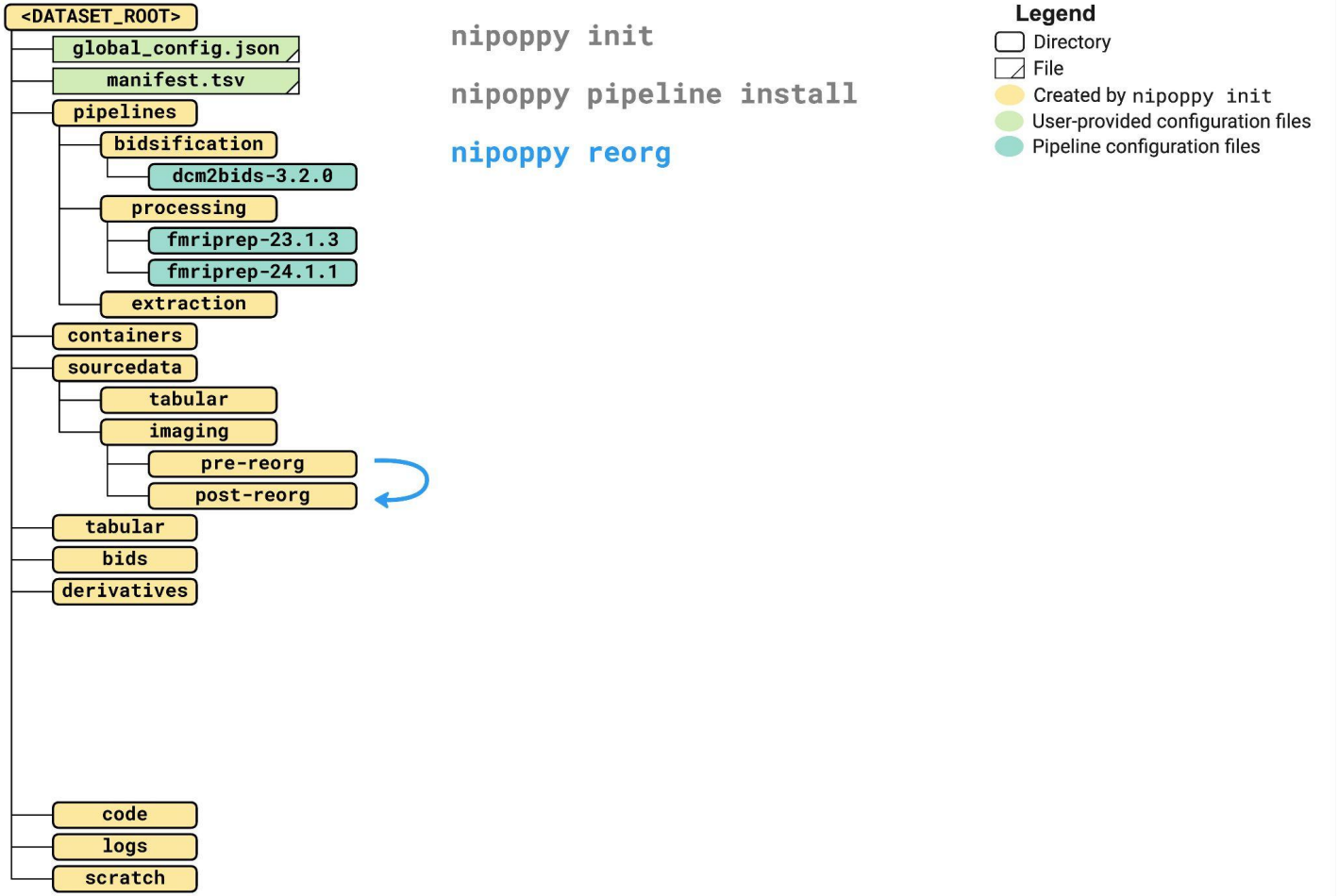
A typical **Nipoppy** workflow



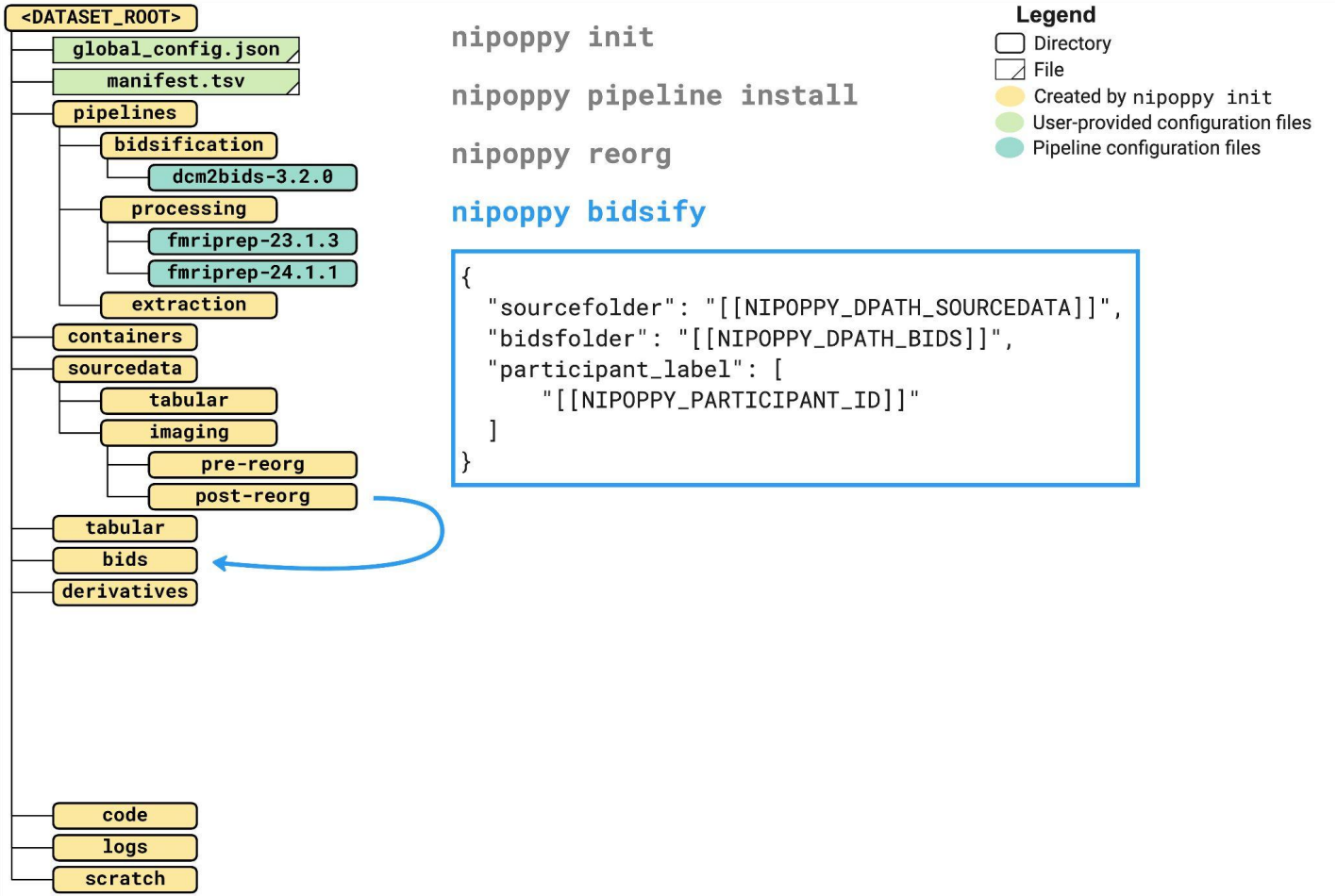
A typical **Nipoppy** workflow



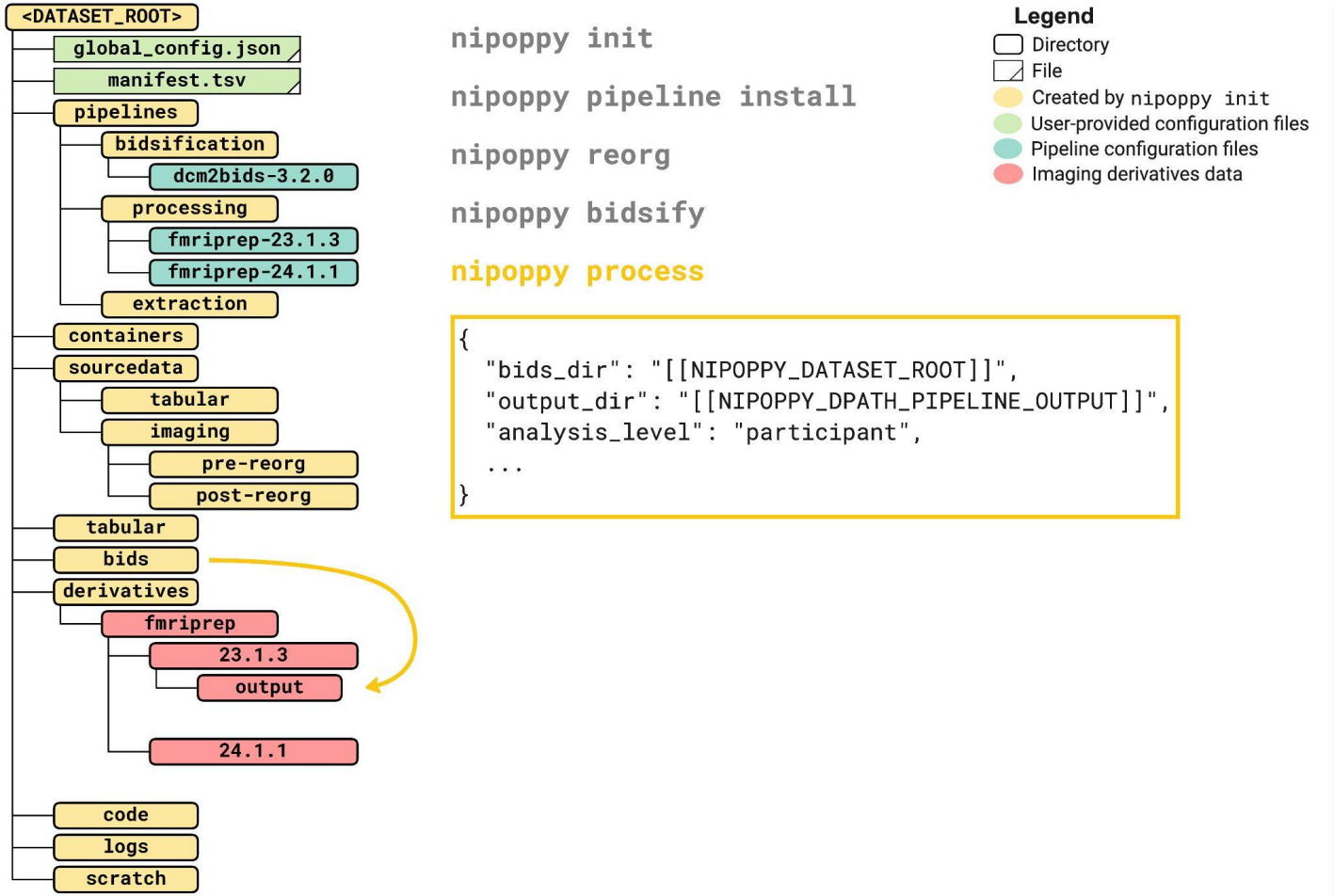
A typical Nipoppy workflow



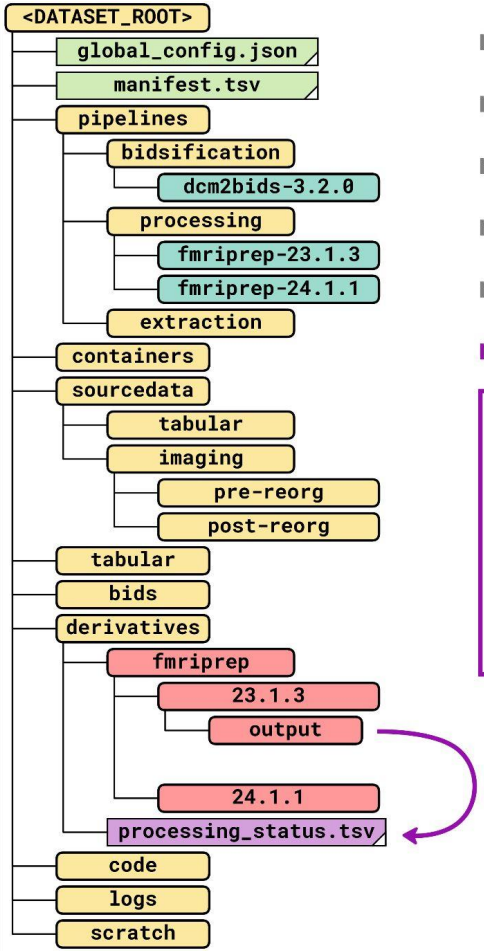
A typical Nipoppy workflow



A typical Nipoppy workflow

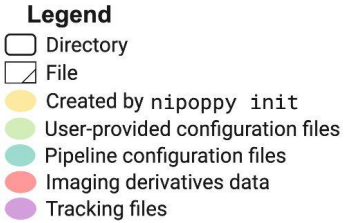


A typical **Nipoppy** workflow

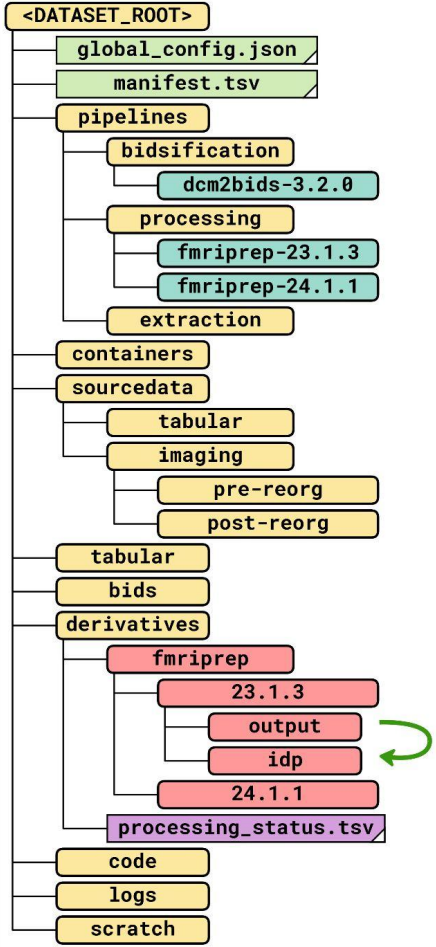



```
nipoppy init

nipoppy pipeline install

nipoppy reorg

nipoppy bidsify

nipoppy process

nipoppy track-processing
```

participant_id	session_id	pipeline	version	status
01	A	freesurfer	6.0.0	SUCCESS
01	A	freesurfer	7.4.1	SUCCESS
02	A	freesurfer	6.0.0	FAIL
02	A	freesurfer	7.4.1	FAIL
02	B	freesurfer	6.0.0	INCOMPLETE
02	B	freesurfer	7.4.1	INCOMPLETE
...	...	...	...	...

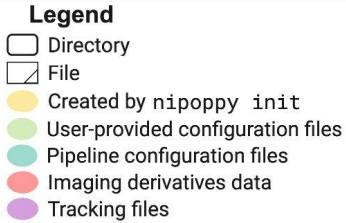


A typical Nipoppy workflow




```
nipoppy init

nipoppy pipeline install

nipoppy reorg

nipoppy bidsify

nipoppy process

nipoppy track-processing

nipoppy extract
```

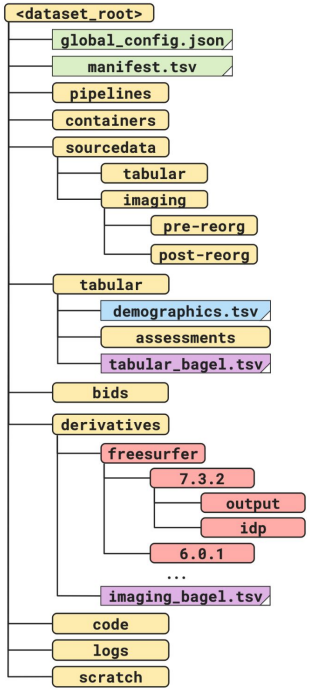
participant_id	session_id	gm_vol	wm_vol	...
001	A	399445	463167	...
002	A	432054	422803	...
002	B	432913	422634	...
...	...	...	...	...



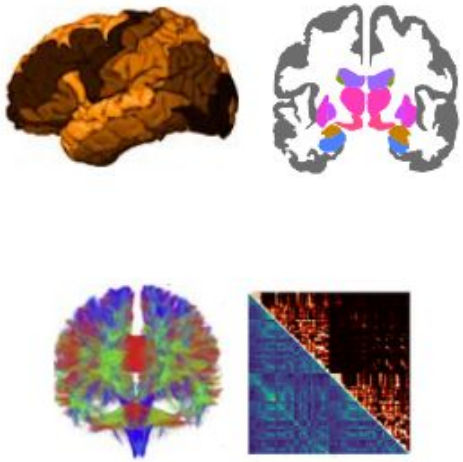
Nipoppy takeaways

Framework combining **existing tools** BIDS BRAIN IMAGING DATA STRUCTURE APPTAINER Boutiques

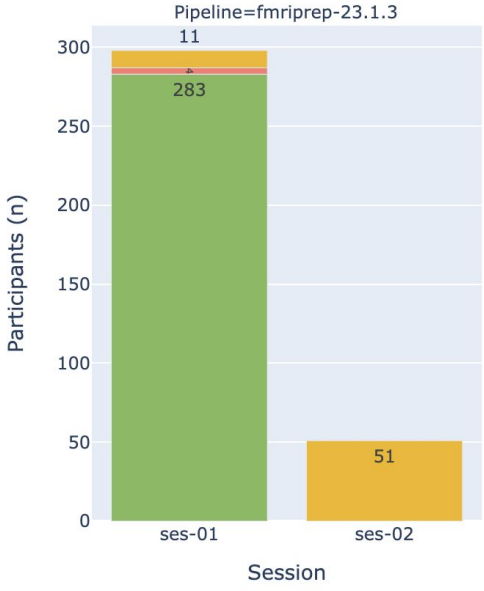
Local data management



Reproducible MRI processing and IDP extraction

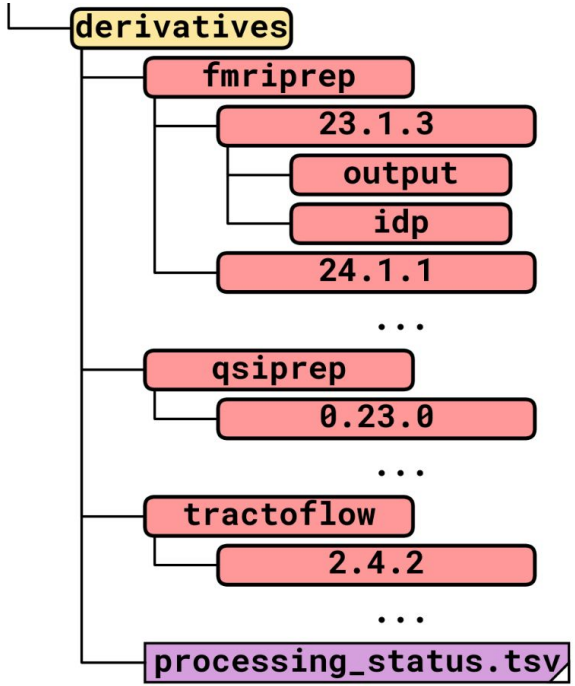


Progress tracking



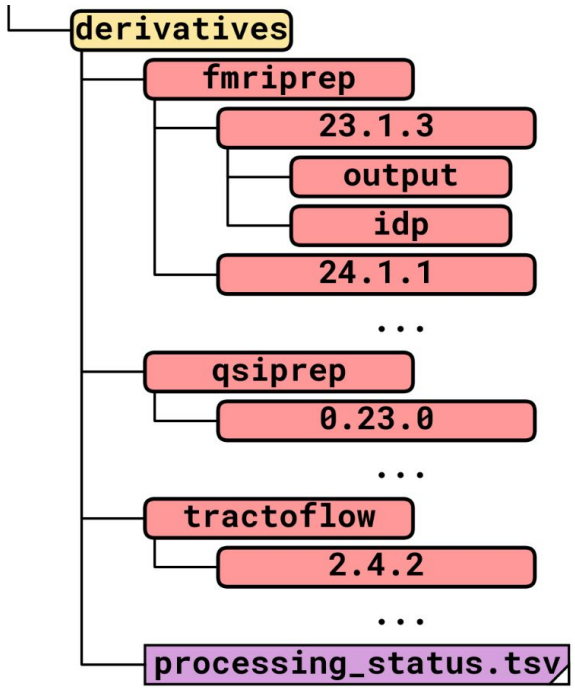
Nipoppy for analytical flexibility studies

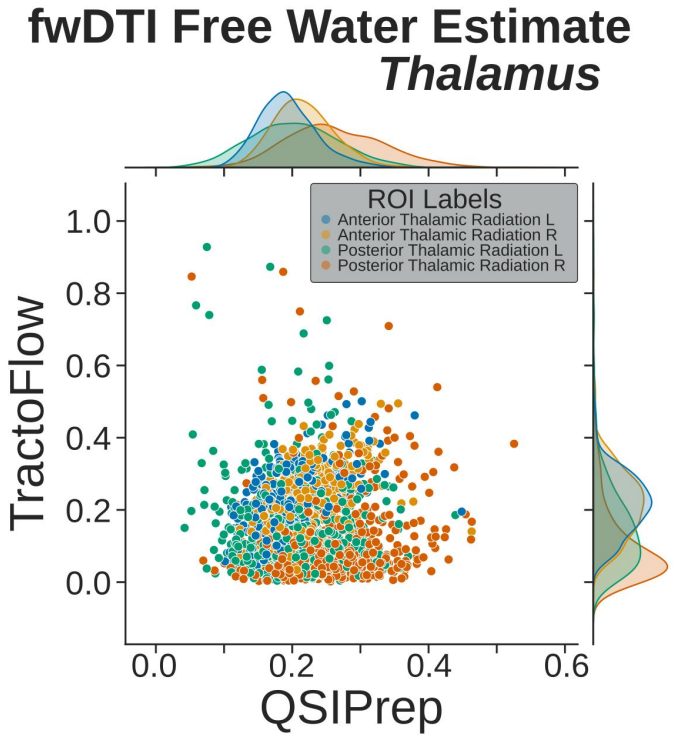
- Running different pipelines/versions
- Runtime parameters stored in JSON files



Nipoppy for analytical flexibility studies

- Running different pipelines/versions
- Runtime parameters stored in JSON files

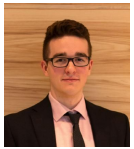




Demo! But before…

Thank you!


Michelle Wang


Mathieu Dugré


Brent McPherson


Nikhil Bhagwat


Arman Jahanpour


Sebastian Urchs


Alyssa Dai


Julia Pfarr


Jean-Baptiste Poline

ORIGAMI Lab

neuro Montreal Neurological Institute-Hospital

CIHR IRSC Canadian Institutes of Health Research / Instituts de recherche en santé du Canada

Québec: Fonds de recherche – Nature et technologies; Fonds de recherche – Santé; Fonds de recherche – Société et culture

PARKINSON QUÉBEC

UNIQUE CENTRE

NIH National Institutes of Health

THE MICHAEL J. FOX FOUNDATION FOR PARKINSON'S RESEARCH

McGill UNIVERSITY

Chan Zuckerberg Initiative

HBHL NEUROHUB

QUEBEC PARKINSON NETWORK

Fondation Brain Canada Foundation