Nipoppy: A framework for the **organization** and **decentralized processing** of neuroimaging-clinical studies

Michelle Wang
Neurodatascience–ORIGAMI Lab (Jean-Baptiste Poline)
McGill University

2024 June 25

OHBM 2024

The truth about many neuroimaging datasets

- They are **messy**
- Data organization/curation and processing is a **time-consuming** and **error-prone** process

How can we organize and process data in a consistent and reproducible manner, to avoid downstream headaches?

Introducing the **Nipoppy** framework

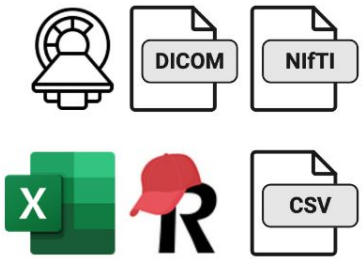
1. **Protocol** for best practices (FAIR principles) in organization/processing of neuroimaging-clinical datasets

Introducing the **Nipoppy** framework

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

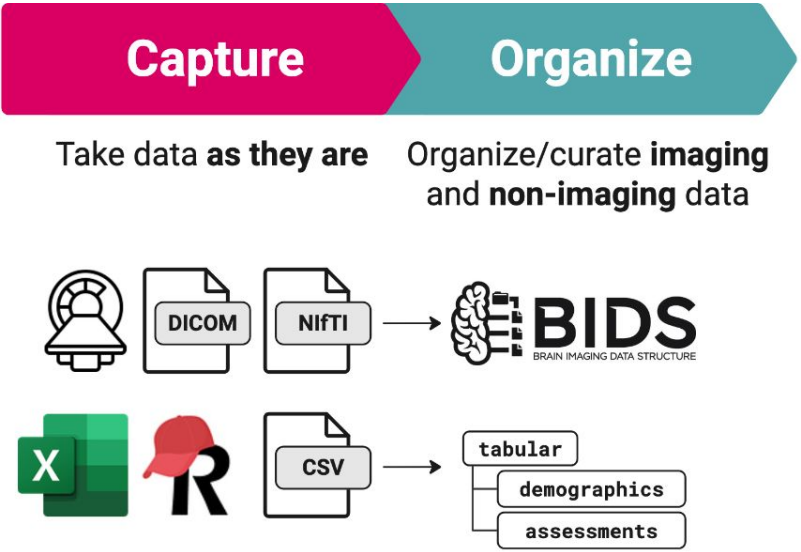
Capture

Take data **as they are**



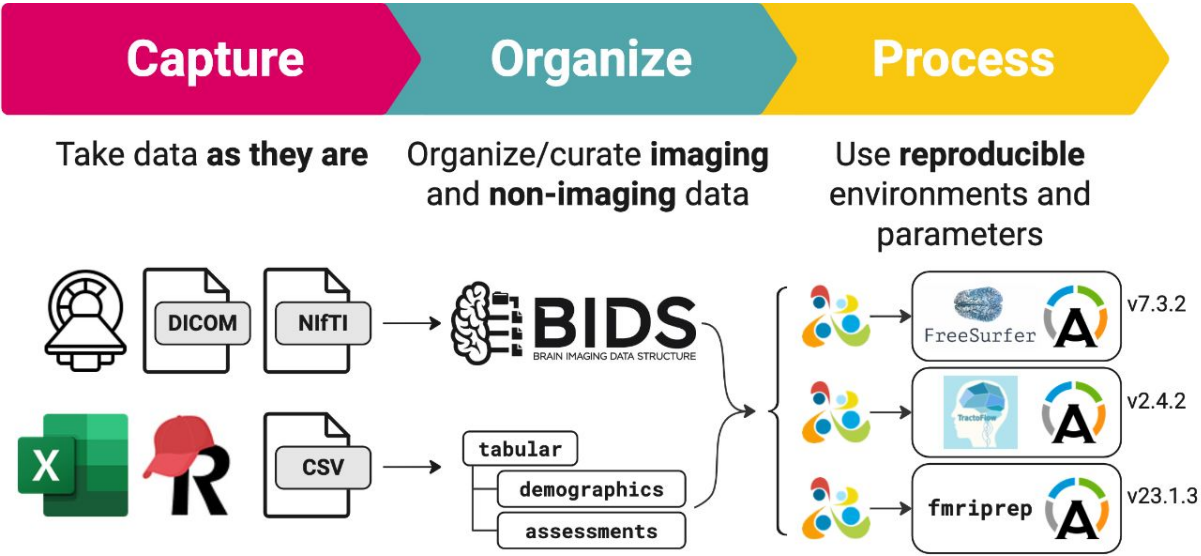
Introducing the **Nipoppy** framework

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



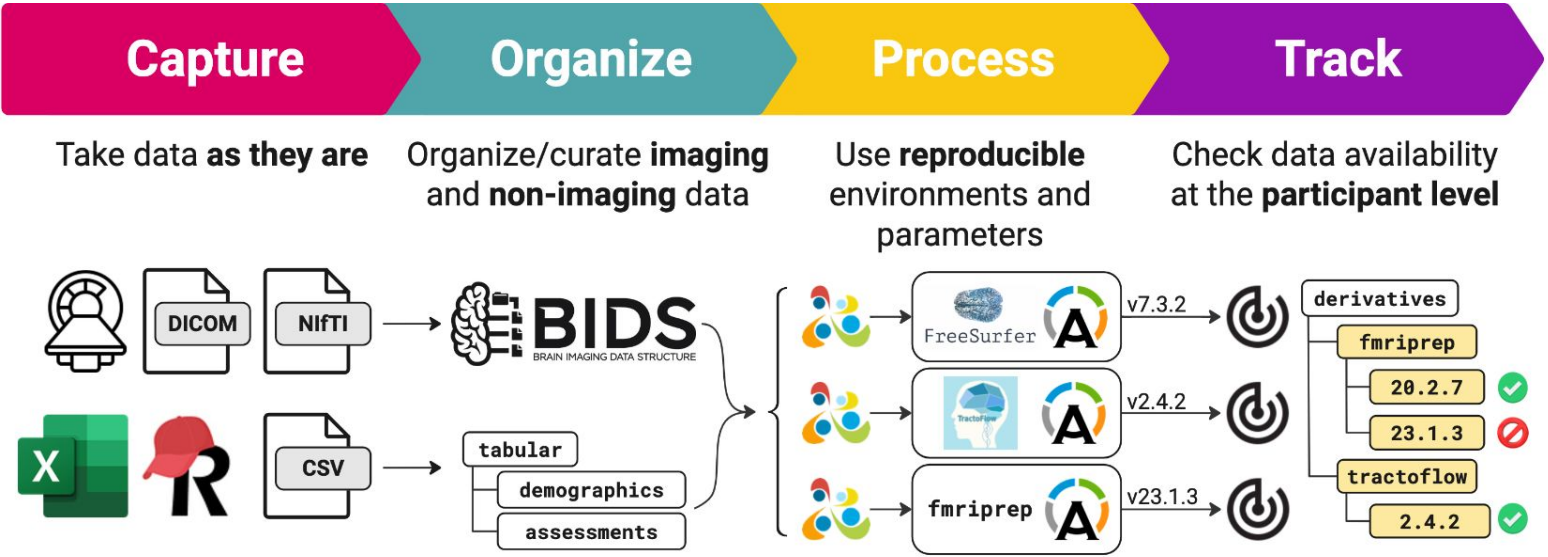
Introducing the **Nipoppy** framework

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



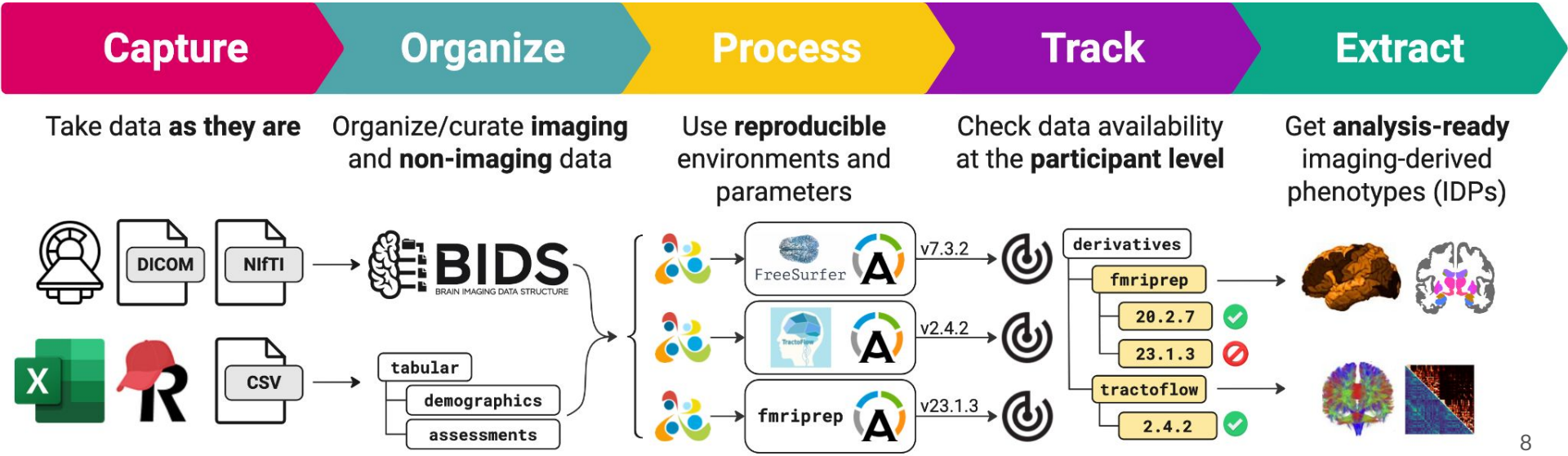
Introducing the **Nipoppy** framework

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



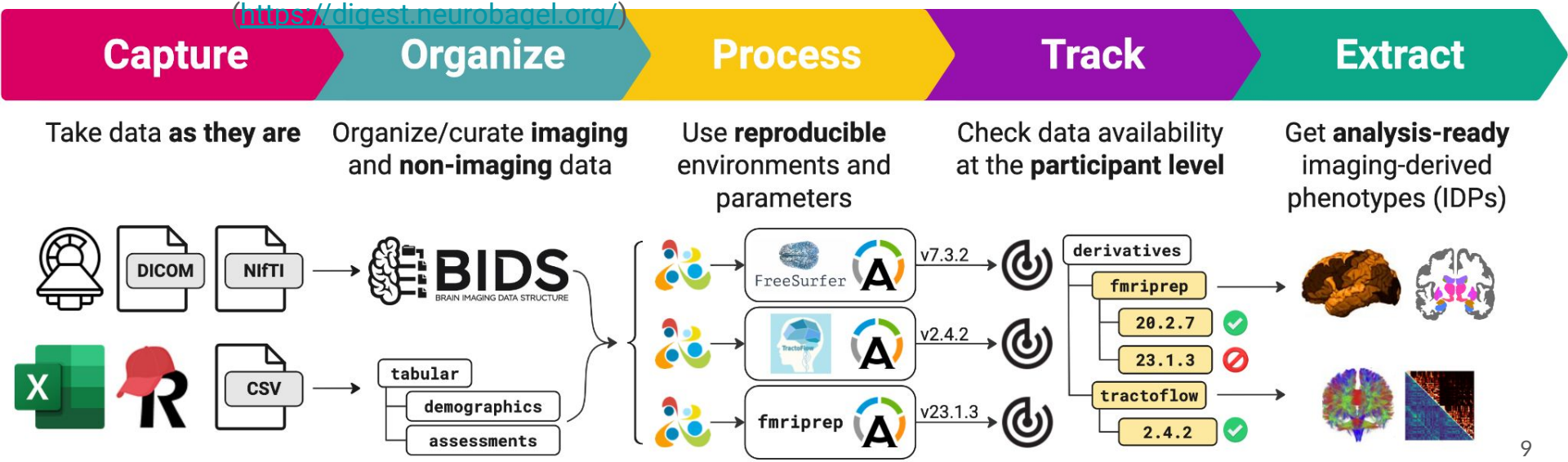
Introducing the **Nipoppy** framework

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



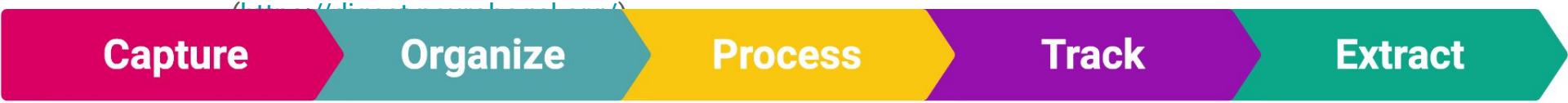
Introducing the **Nipoppy** framework

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

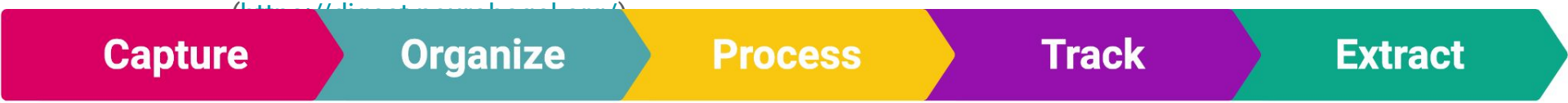
1. **Protocol** for best practices (FAIR principles) in organization/processing of neuroimaging-clinical datasets
   a. From raw scanner data to extracted imaging-derived phenotypes (IDPs)
   b. Visualizing tabular data/metadata on a dashboard



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

Introducing the **Nipoppy** framework

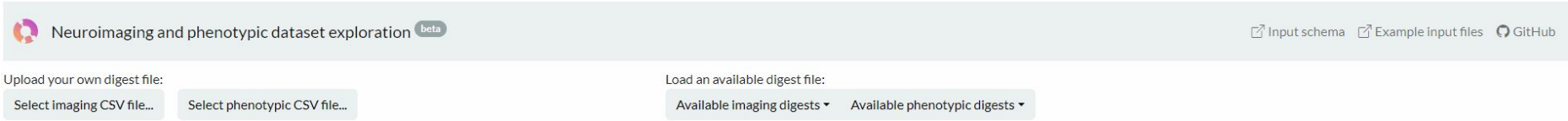
1. **Protocol** for best practices (FAIR principles) in organization/processing of neuroimaging-clinical datasets
   a. From raw scanner data to extracted imaging-derived phenotypes (IDPs)
   b. Visualizing tabular data/metadata on a dashboard



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

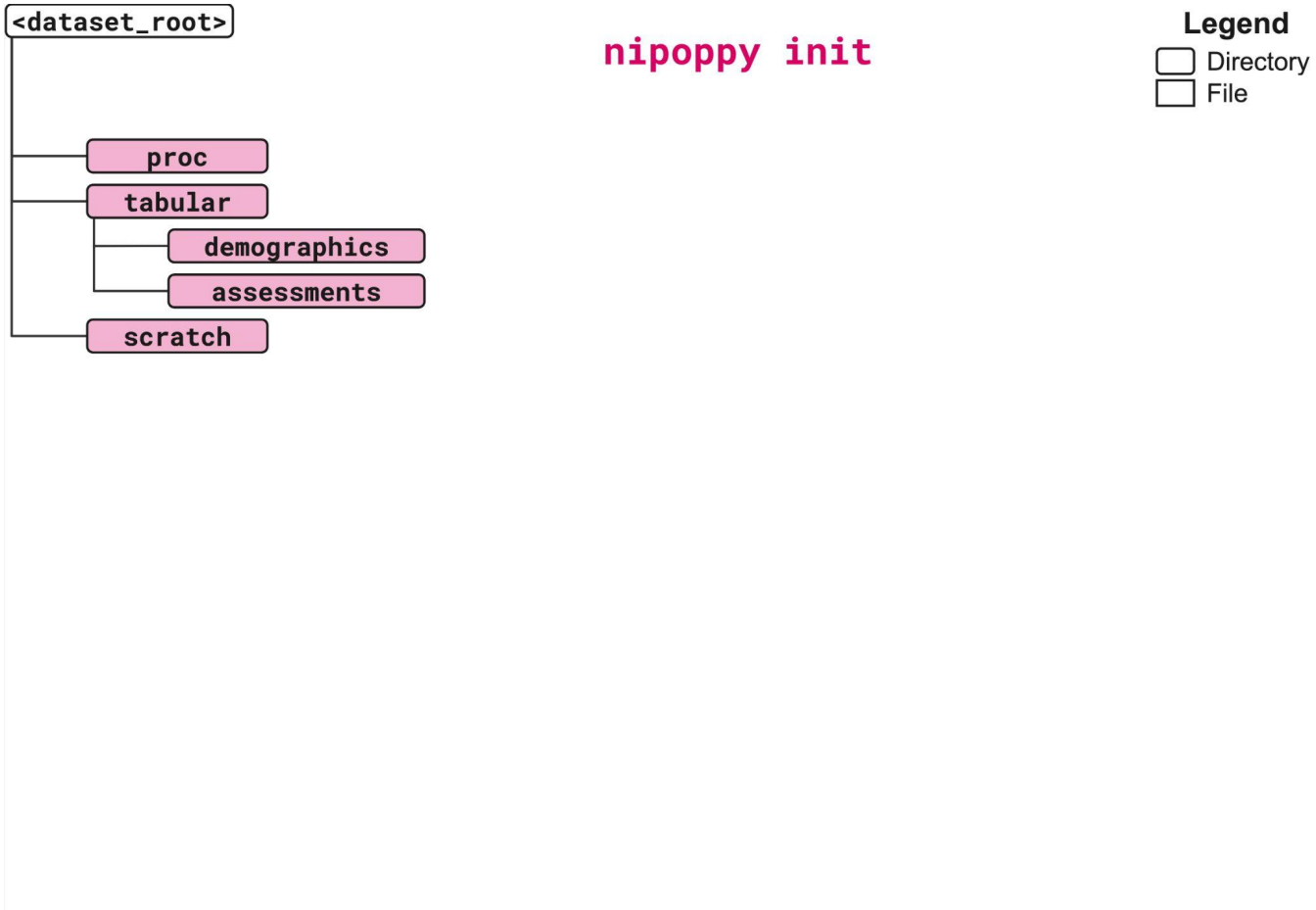
3. Software package with **tools** to help work with this framework

https://digest.neurobagel.org/

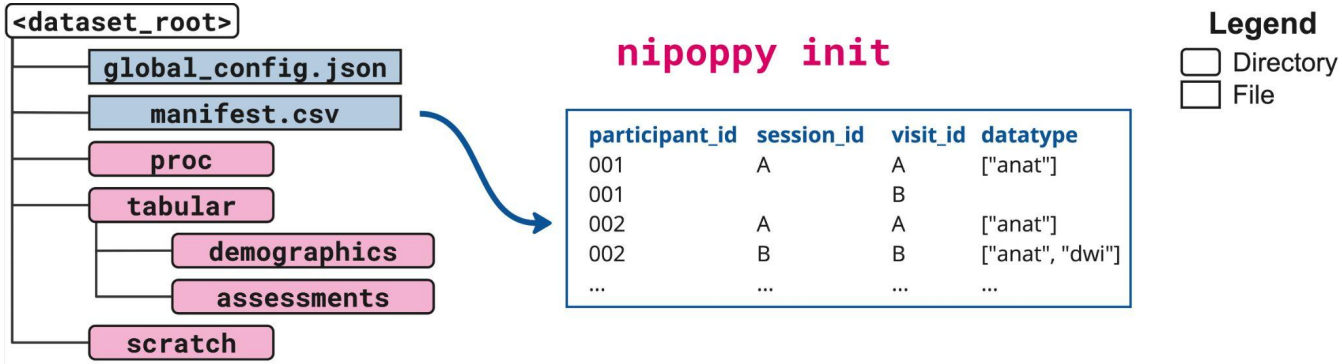


A typical **Nipoppy** workflow

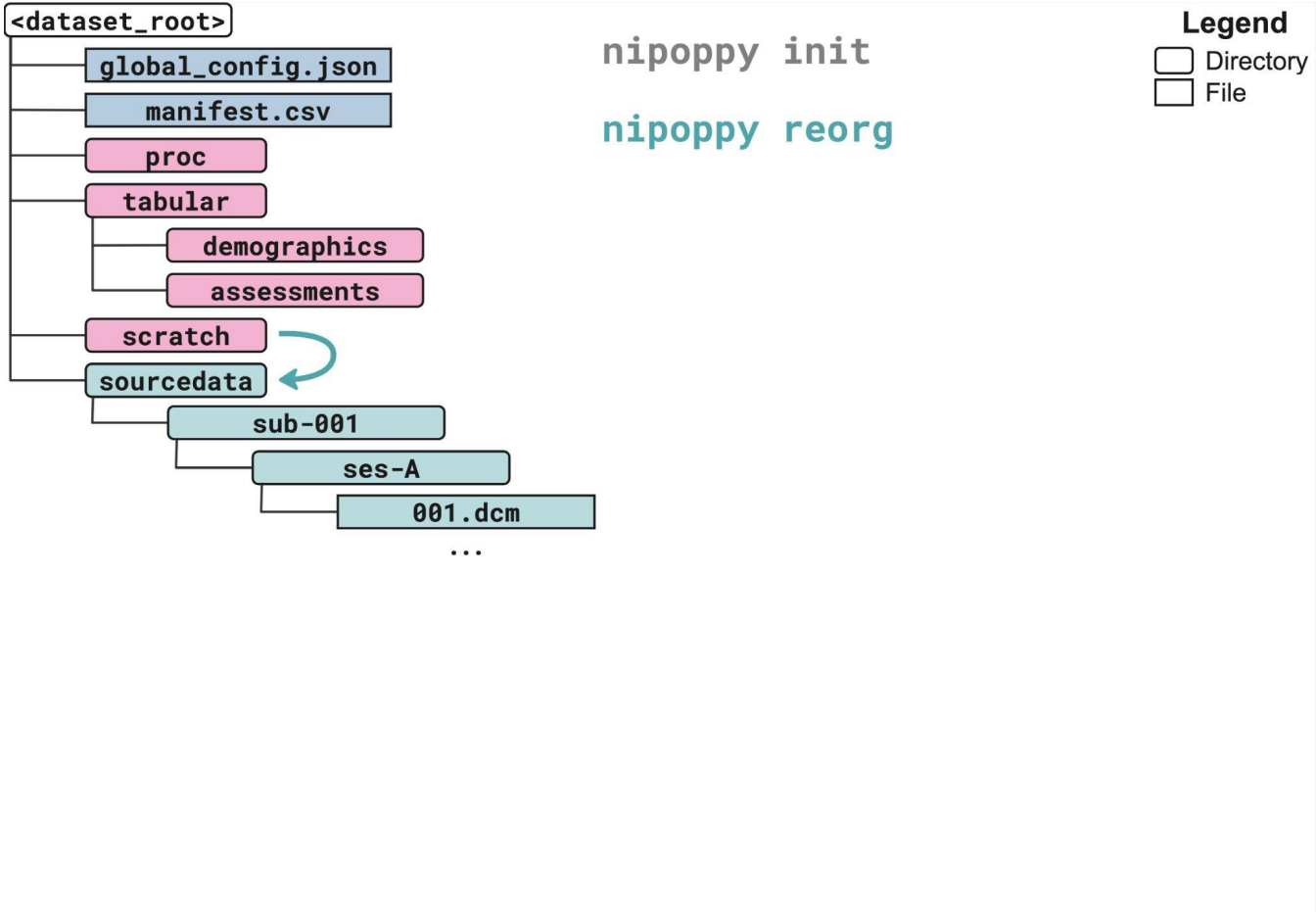
A typical **Nipoppy** workflow



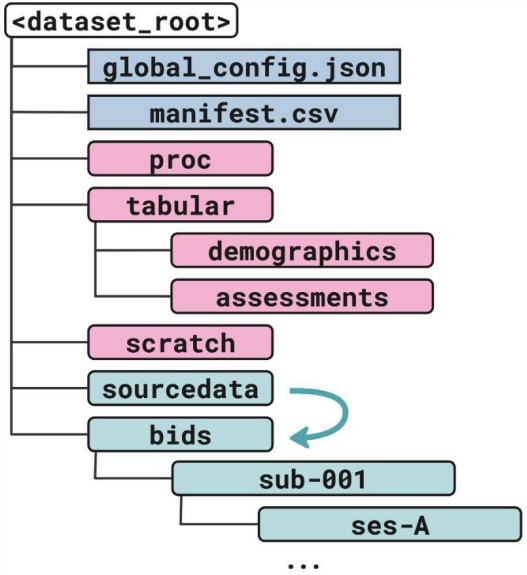
A typical **Nipoppy** workflow



A typical **Nipoppy** workflow



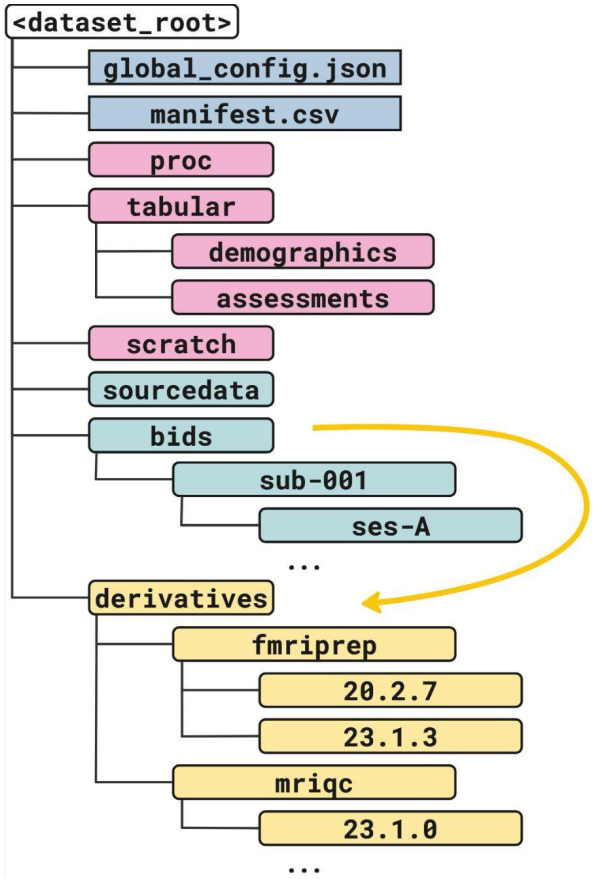
A typical Nipoppy workflow




```
nipoppy init

nipoppy reorg

nipoppy bidsify
```

```
{
    "sourcefolder": "[[NIPOPPY_DPATH_SOURCEDATA]]",
    "bidsfolder": "[[NIPOPPY_DPATH_BIDS]]",
    "participant_label": [
        "[[NIPOPPY_PARTICIPANT_ID]]"
    ]
}
```

A typical Nipoppy workflow



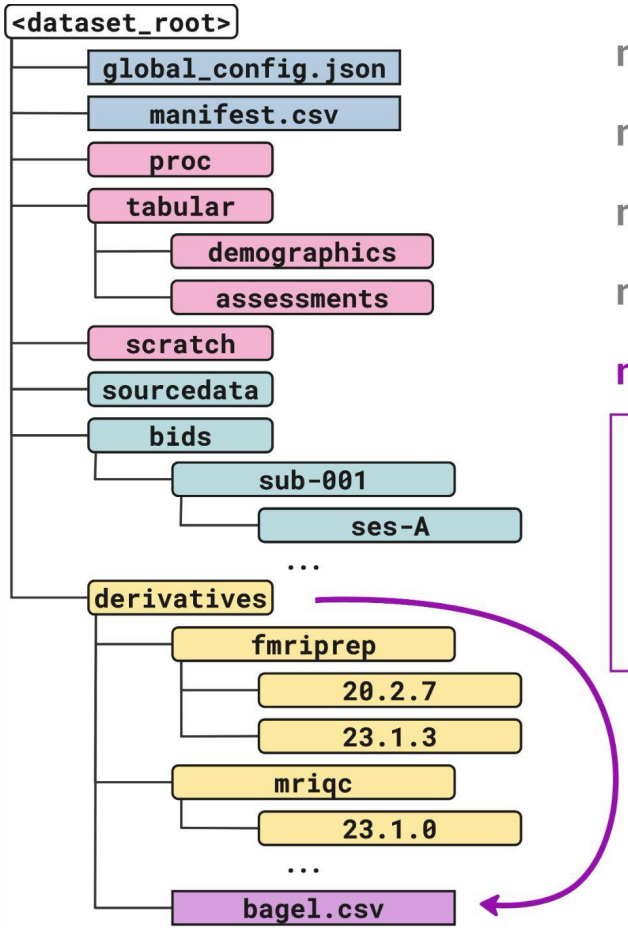

```
nipoppy init

nipoppy reorg

nipoppy bidsify

nipoppy run
```

```
{
    "bids_dir": "[[NIPOPPY_DATASET_ROOT]]",
    "output_dir": "[[NIPOPPY_DPATH_PIPELINE_OUTPUT]]",
    "analysis_level": "participant",
    ...
}
```

Legend

- Directory
- File

A typical **Nipoppy** workflow

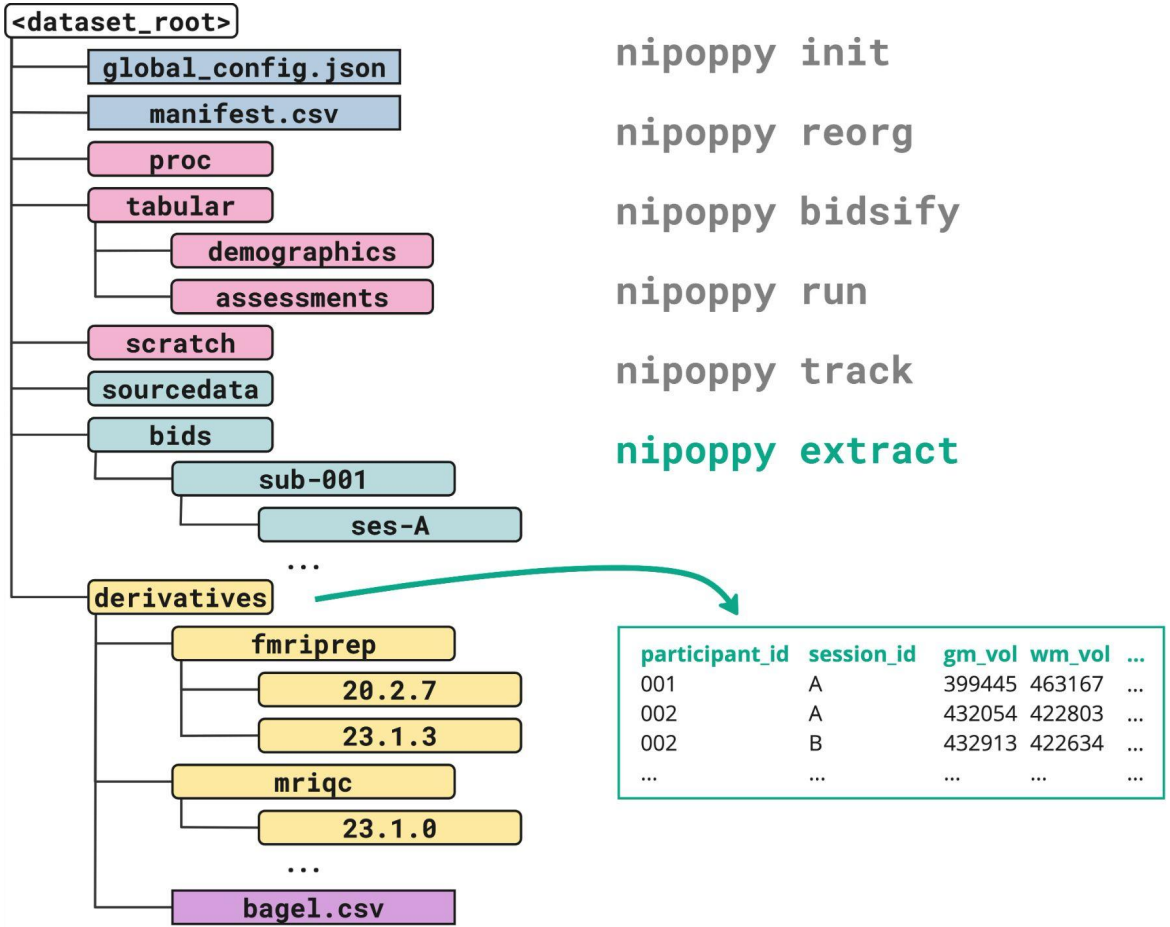



```
nipoppy init

nipoppy reorg

nipoppy bidsify

nipoppy run
```

nipoppy track

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 takeaways

Framework combining **existing tools**
Boutiques

- **Standardized organization** and workflow for both imaging and tabular data
- **Flexible** and **reproducible** processing pipelines
- Pipeline output **tracking** and imaging-derived phenotype (IDP) **extraction**
- Built-in support for **ongoing studies**
- Local, **decentralized** setup → facilitate multisite studies (see also: Neurobagel)

Thank you!

- Contact: michelle.wang6@mail.mcgill.ca
- Poster **#2256** (June 26-27, 13:45-15:45)
- Exhibitor booth **#103**
- https://github.com/neurodatascience/nipoppy
- https://nipoppy.readthedocs.io/en/latest/
- Related project: https://neurobagel.org/

Nikhil Bhagwat

Brent McPherson

Rémi Gau

Alyssa Dai

Sebastian Urchs

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