

Reproducible Project Management

Dr. Anwar Said

Institute for Software Integrated Systems,
Vanderbilt University, TN

Reproducible Project Management



Open-sourcing



Documentation

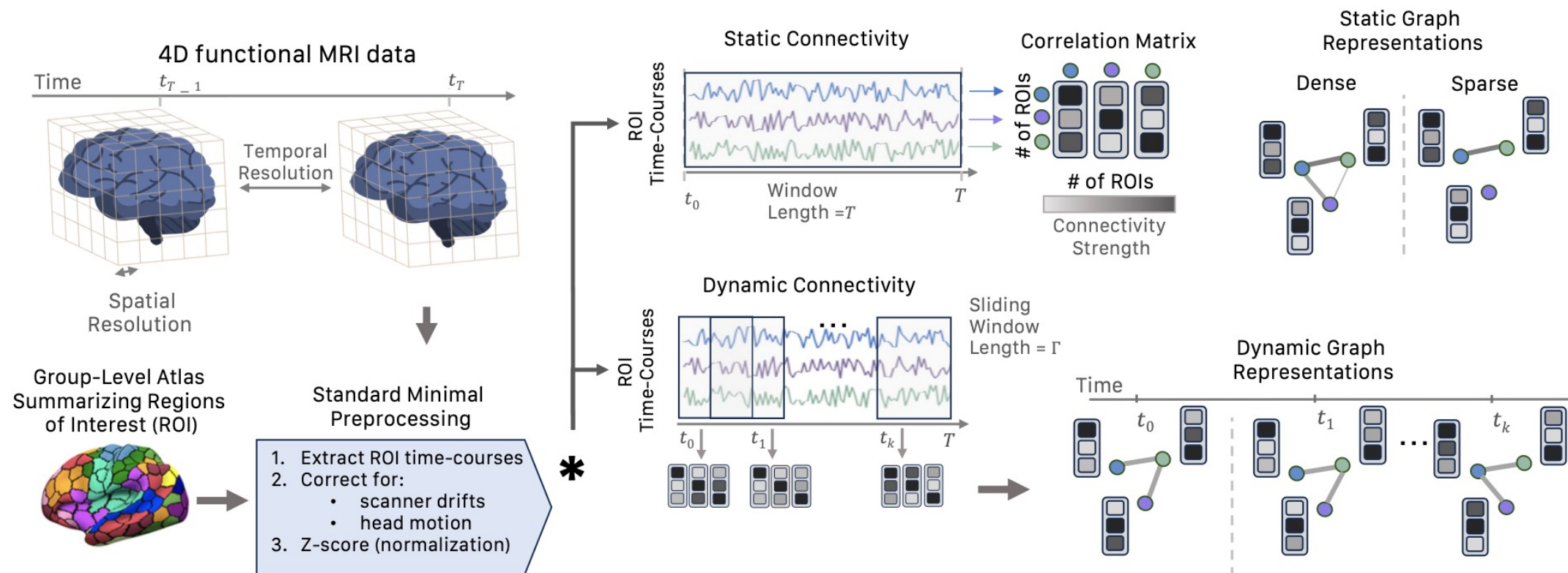


Packaging




Version control

NeuroGraph: Example



NeuroGraph Docs

 NeuroGraph

latest

Search docs

NEUROGRAPH:
Demographics
Mental States
Cognitive Traits

INSTALLATIONS:
Installation

GET STARTED:
Introduction by Example
Preprocessing Examples
Preprocessing Human Connectome Project (HCP1200) Dataset

LOADING BENCHMARKS:
Load Benchmark Datasets

PREPROCESSING:
NeuroGraph Preprocessing Functionalities

 / Indices and tables

NeuroGraph:

- [Demographics](#)
- [Mental States](#)
- [Cognitive Traits](#)

Installations:

- [Installation](#)

GET STARTED:

- [Introduction by Example](#)
 - [Loading Benchmark datasets](#)
- [Preprocessing Examples](#)
- [Preprocessing Human Connectome Project \(HCP1200\) Dataset](#)
 - [Download and preprocess static datasets](#)
 - [Download and preprocess dynamic datasets](#)

Loading Benchmarks:



Search projects

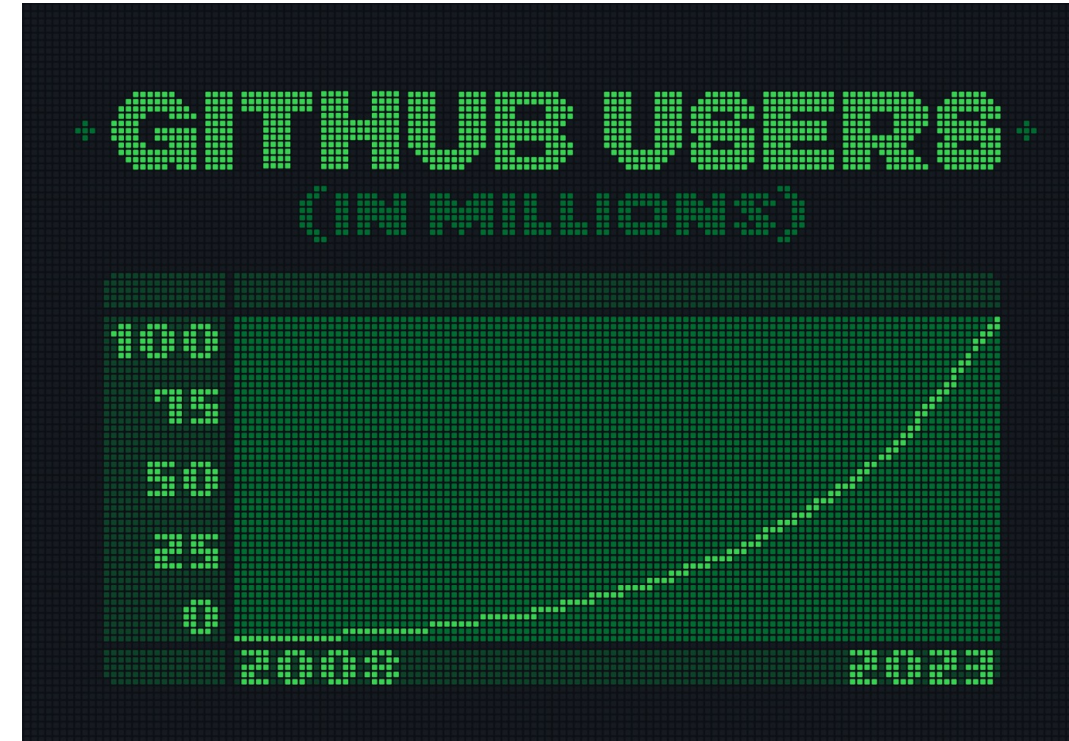
NeuroGraph 2.4.0

```
pip install NeuroGraph
```



Open-sourcing with GitHub

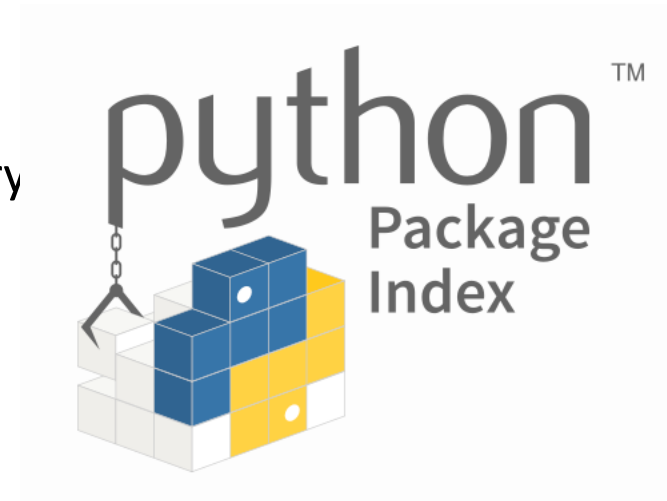
- A website that hosts git repositories on a remote server
- Facilitates the sharing of codebases among teams by providing a GUI to easily fork or clone repos to a local machine
- Why?
- Collaboration and contribution
- Community and visibility



<https://github.blog/2023-01-25-100-million-developers-and-counting/>

Packaging Python projects

- Python Packaging Index (PyPi)
- Repository of software for the Python programming language
- Why?
- Official package distribution and discovery repository
- Standardized package distribution format



506,760 projects

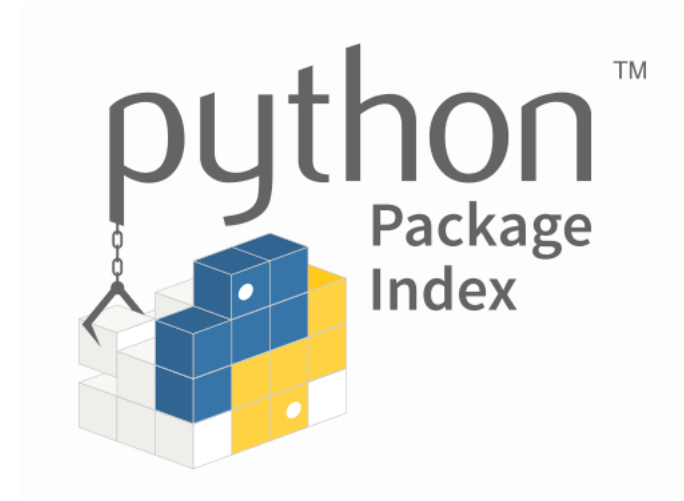
5,268,097 releases

10,047,312 files

775,126 users

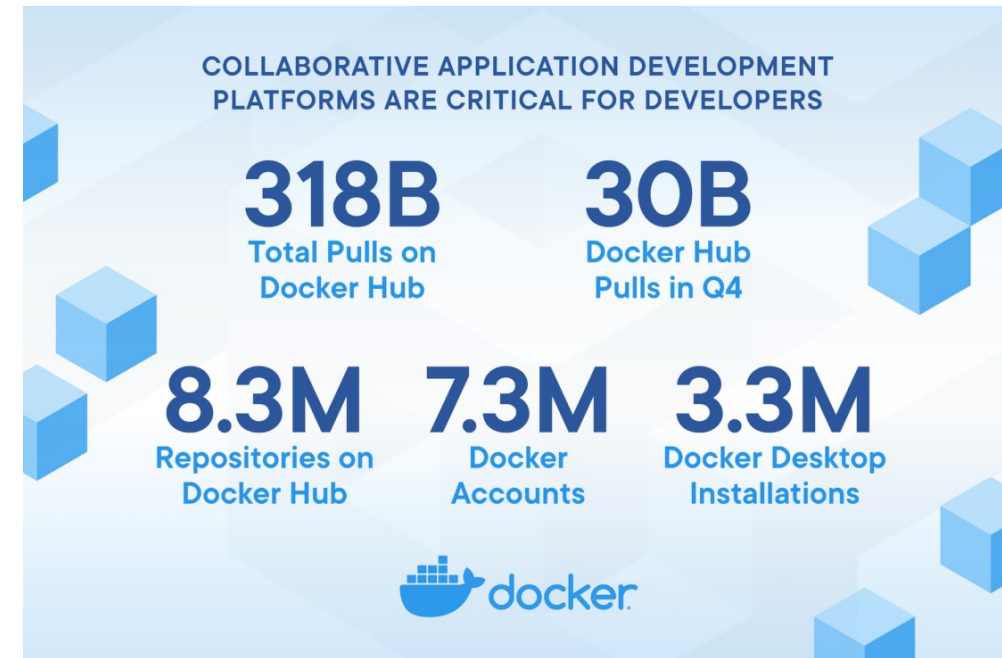
Packaging Python projects

- Python Packaging Index (PyPi)
- Repository of software for the Python programming language
- **Why?**
- Official package distribution and discovery repository
- Standardized package distribution format
- **Walk you through:**
 - How to package a Python project
 - How to add the necessary files and structure
 - Build and upload the package to PyPi
 - How to install and use



Containerize Python Project

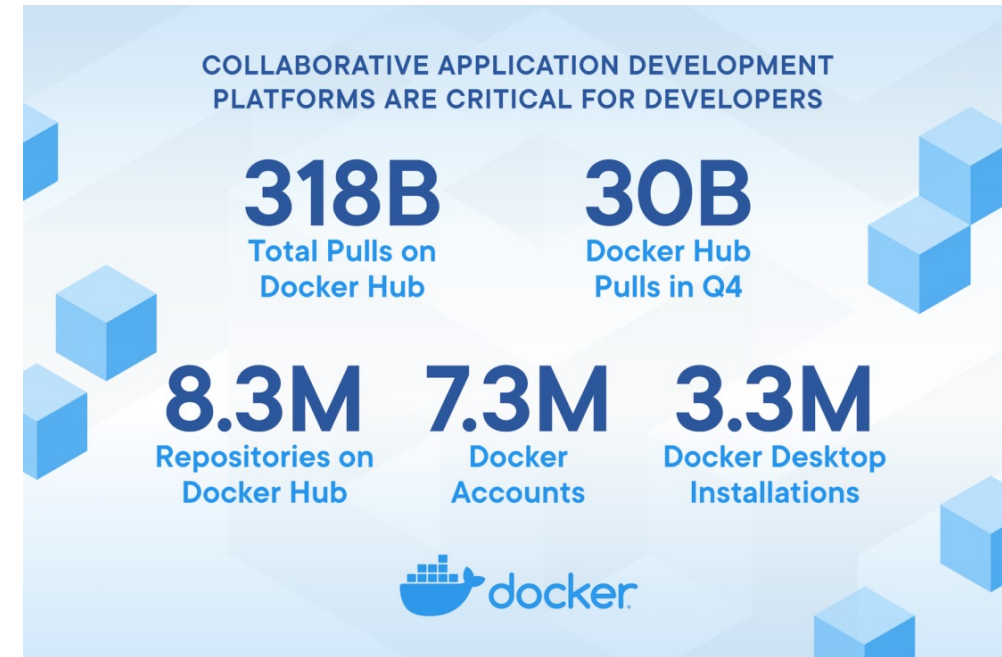
- Containerize means to create a Docker container for your project
- It encapsulate the application dependencies and runtime environment
- **Why?**
 - Consistent environment
 - Easier dependency management
 - Faster onboarding
 - Simplified deployment and many others



<https://www.docker.com/blog/docker-index-shows-continued-massive-developer-adoption-and-activity-to-build-and-share-apps-with-docker/>

Containerize Python Project


- Containerize means to create a Docker container for your project
- It encapsulate the application dependencies and runtime environment
- **Why?**
 - Consistent environment
 - Easier dependency management
 - Faster onboarding
 - Simplified deployment and many others
- Walk you through:
 - How to add the necessary files and structure
 - Build the Docker and upload to the Docker Hub
 - How to download and use



<https://www.docker.com/blog/docker-index-shows-continued-massive-developer-adoption-and-activity-to-build-and-share-apps-with-docker/>

Documentation with Read the Docs

- Simplifies software documentation
- Automatically perform building, versioning and hosting docs
- Walk you through:
 - How to add the necessary files and structure
 - Build and publish the documentation




latest

INSTALL PYG
Installation

GET STARTED
Introduction by Example
Colab Notebooks and Video Tutorials

TUTORIALS
Design of Graph Neural Networks
Working with Graph Datasets
Use-Cases & Applications
Multi-GPU Training

ADVANCED CONCEPTS
Advanced Mini-Batching
Memory-Efficient Aggregations

 Read the Docs v: latest ▾

NeuroGraphDataset

```
class NeuroGraphDataset ( root: str, name: str, transform: Optional[Callable] = None,
pre_transform: Optional[Callable] = None, pre_filter: Optional[Callable] = None, force_reload:
bool = False ) \[source\]
```

Bases: `InMemoryDataset`

The NeuroGraph benchmark datasets from the “[NeuroGraph: Benchmarks for Graph Machine Learning in Brain Connectomics](#)” paper. `NeuroGraphDataset` holds a collection of five neuroimaging graph learning datasets that span multiple categories of demographics, mental states, and cognitive traits. See the [documentation](#) and the [Github](#) for more details.

Dataset	#Graphs	Task
<code>HCPActivity</code>	7,443	Graph Classification
<code>HCPGender</code>	1,078	Graph Classification
<code>HCPAge</code>	1,065	Graph Classification
<code>HCPFI</code>	1,071	Graph Regression
<code>HCPWM</code>	1,078	Graph Regression

PARAMETERS:

- `root` (`str`) – Root directory where the dataset should be saved.



<https://github.com/brainhack-vandy/NeuroPreprocessing>