Developing an fMRI workflow – lessons learned on integration and reproducibility

Josh McGeown[1] and Maxime Rio[2,3]

[1]Mātai, [2]NeSI, [3]NIWA

j.mcgeown@matai.org.nz, maxime.rio@nesi.org.nz

ABSTRACT / INTRODUCTION (Up to 200 words)

Advanced MRI sequences provide invaluable data with high temporal and spatial resolution to understand how the brain can be affected by a particular condition. Before deriving scientific conclusions from a dataset collected on a cohort of subjects, one need to process many individual images in a series of preprocessing steps before aggregating the results across subjects. Opensource toolboxes are available to help with part of this process. However, it can be challenging to chain them and scale the computations for large datasets.

In this talk, we will present the work we did to coordinate these tools in a coherent and reproducible workflow using Snakemake. We will discuss the delicate art of gluing tools with their own set of assumptions, praise the beauty of community owned open data formats and digress on our adventures in scaling up computations on an HPC while keeping portability.

ABOUT THE AUTHOR(S)

Josh McGeown is a neuroscientist specializing in traumatic brain injury. Josh's current work leverages multimodal neuroimaging, wearable sensors, and AI/ML in an effort to create a complete picture of the brain's state after exposure to brain injury and how changes in this state relate to recovery patterns.

Maxime Rio is a data science engineer and data scientist at NeSI and NIWA. He enjoys helping researchers to analyse their data, from visualisation to probabilistic modelling.