

CONSIDERATIONS AND CHALLENGES FOR FAIR IN LARGE AND HETEROGENEOUS SETS OF PERSONALIZED LARGE-SCALE MODELS

KELLY SHEN, PHD

INCF ASSEMBLY 2021, SESSION 3

“FAIR APPROACHES FOR COMPUTATIONAL NEUROSCIENCE”


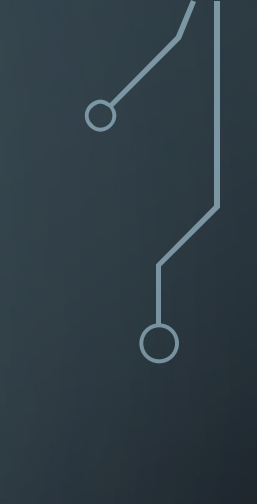
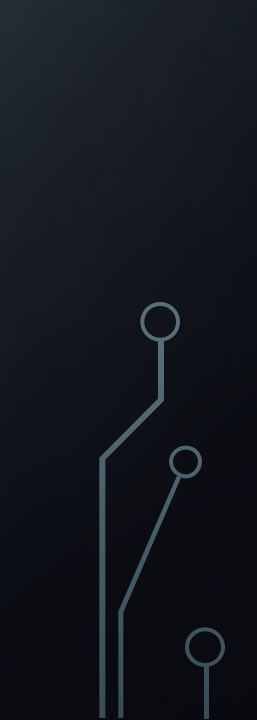
APRIL 20, 2021

Baycrest

Rotman Research Institute



OUTLINE

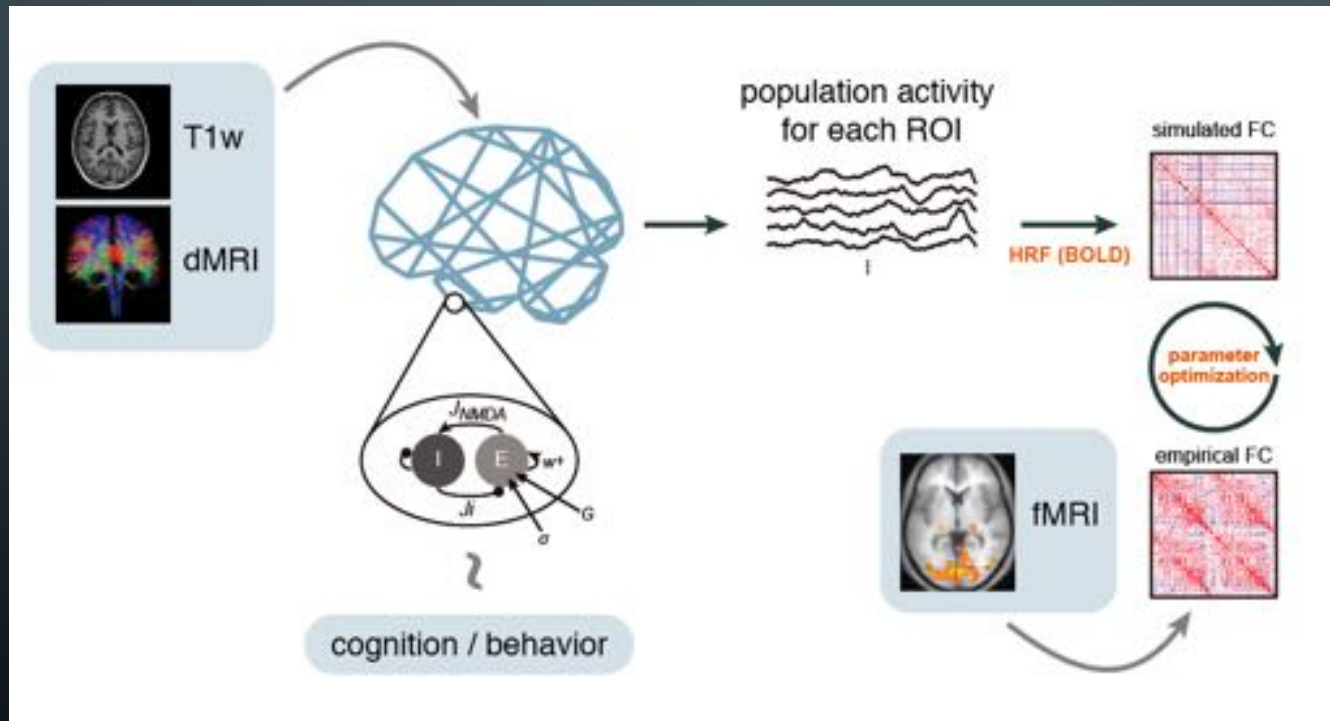
1. Data-driven personalized models: Workflow for big data
 2. Challenges to FAIR in personalized models
 3. FAIR practices throughout the workflow
 4. Improving FAIR practices & scaling up to the cloud
 5. Remaining challenges to FAIR in personalized models
- 
- 
- 

PERSONALIZED MODELS IN THE AGE OF **BIG DATA**

- Opportunity for large and heterogeneous sets of models using data from large data sharing initiatives; for example:
 - Human Connectome Project (HCP)
 - UK Biobank (UKBB)
 - **Alzheimer's Disease Neuroimaging Initiative (ADNI)**
 - **Parkinson's Progressive Markers Initiative (PPMI)**
 - **Cambridge Centre for Ageing and Neuroscience dataset (Cam-CAN)**

PERSONALIZED MODELS IN THE VIRTUAL BRAIN

- Large-scale network models can be personalized to an individual using any number of neuroimaging modalities



Example TVB Workflow

CHALLENGES TO FAIR IN PERSONALIZED MODELS

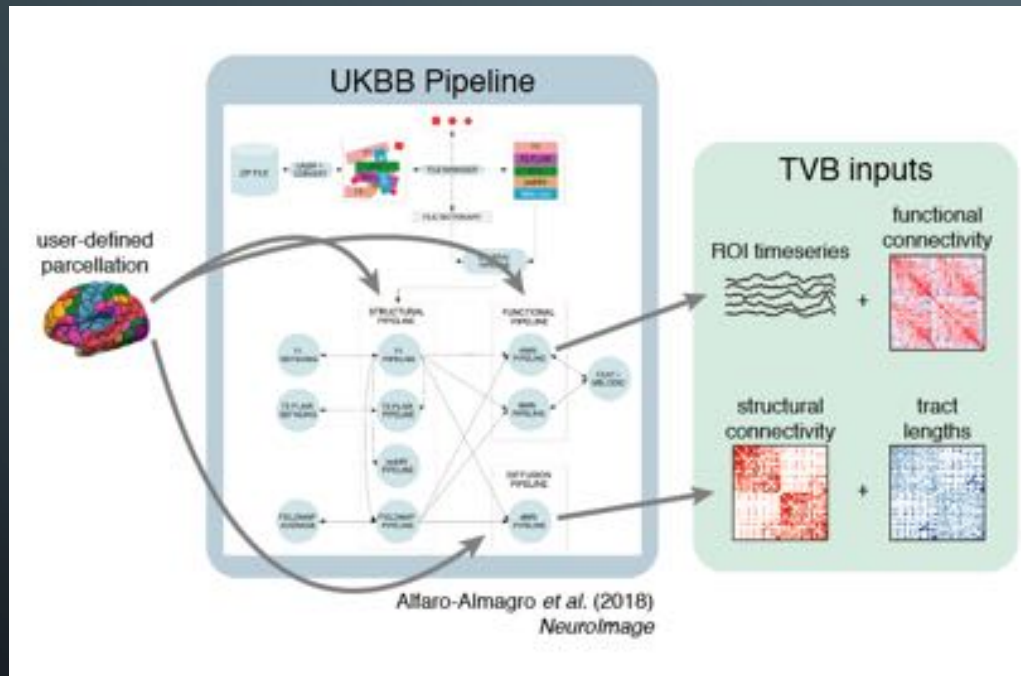
- Long pipeline with potentially many contributors from various disciplines
- Empirical data processing choices can potentially have profound effects on model behavior
- Quality control of processed empirical data is essential
- Size of a “complete” dataset can potentially be very large (both empirical and model derivatives)
- Legal and ethical requirements for the protection of personal data

FAIR PRACTICES START WITH OUR DATA MANAGEMENT PLAN

- The data management plan outlines:
 1. Data collection procedures or data origin
 2. Documentation and metadata requirements
 3. Data storage plan (including back-ups & long-term preservation)
 4. Sharing and reuse plans
 5. Responsible parties & resource requirements (e.g., hardware, \$)
 6. Compliance with ethical and legal requirements

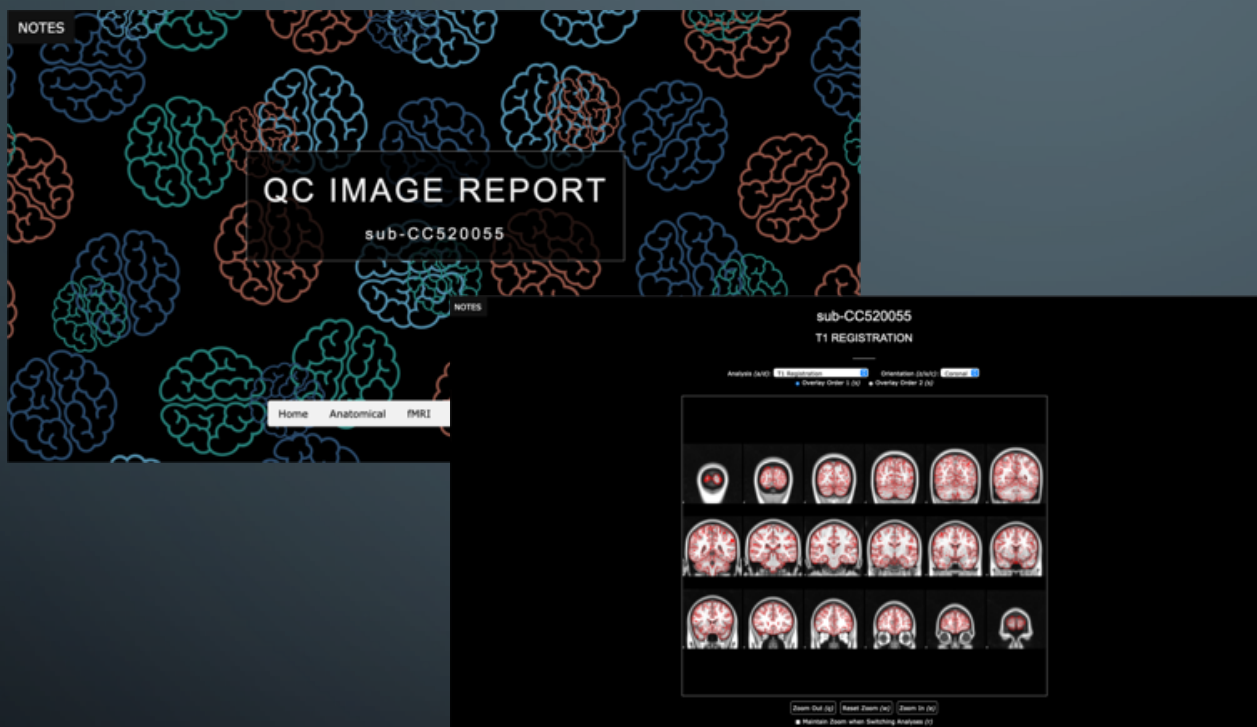
also see <https://portagenetwork.ca> for resources on DMPs

FAIR IN EMPIRICAL DATA PROCESSING: THE TVB-UKBB NEUROIMAGE PROCESSING PIPELINE



- “standardized” processing pipeline for various MRI modalities
 - robust, automated & high-throughput
- BIDS-ified empirical outputs directly interface with TheVirtualBrain

FAIR PRACTICES START WITH EMPIRICAL DATA PROCESSING: THE TVB-UKBB NEUROIMAGE PROCESSING PIPELINE

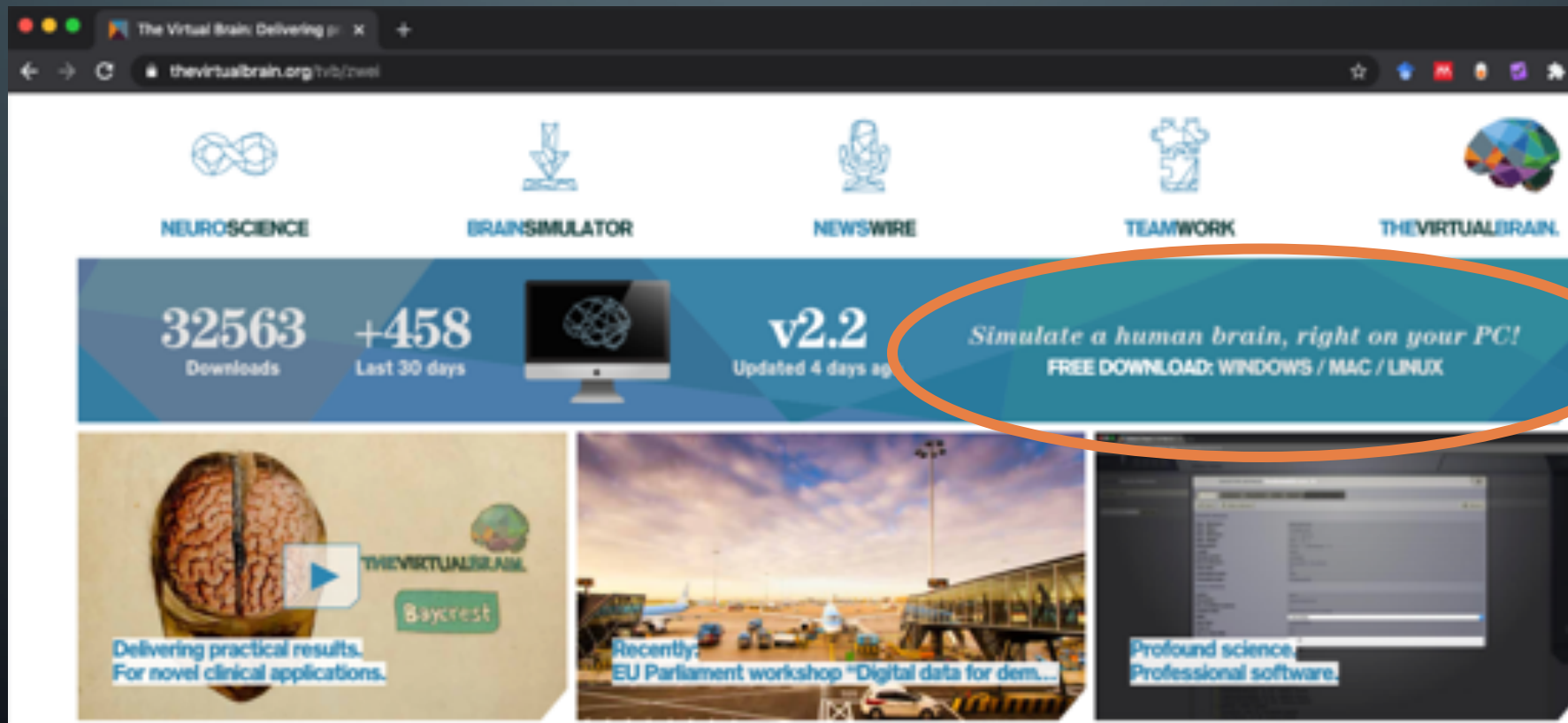


- detailed Quality Control (QC) reports with standardized annotations

See our poster at INCF Assembly 2021:
Noah Frazier-Logue *et al.*

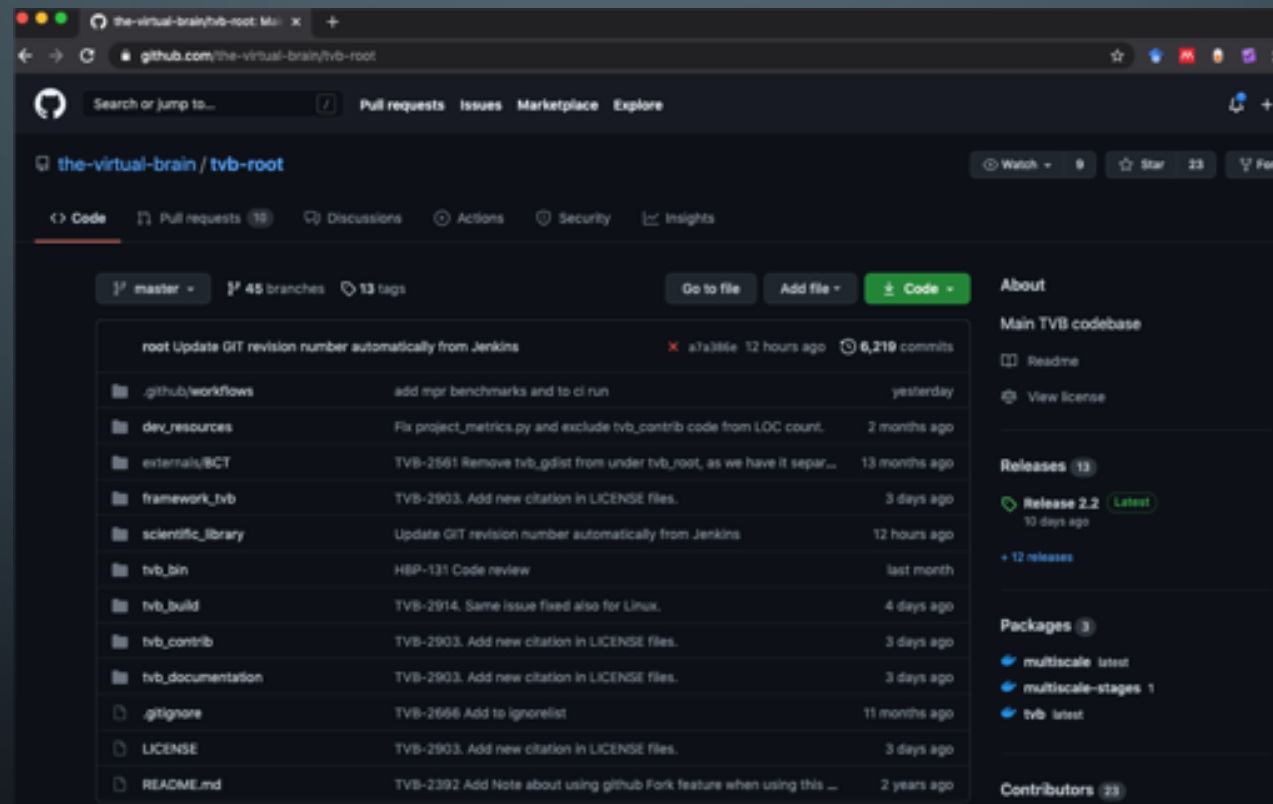
FAIR PRACTICES & THEVIRTUALBRAIN

- available for Linux, Windows and Mac OS (thevirtualbrain.org)

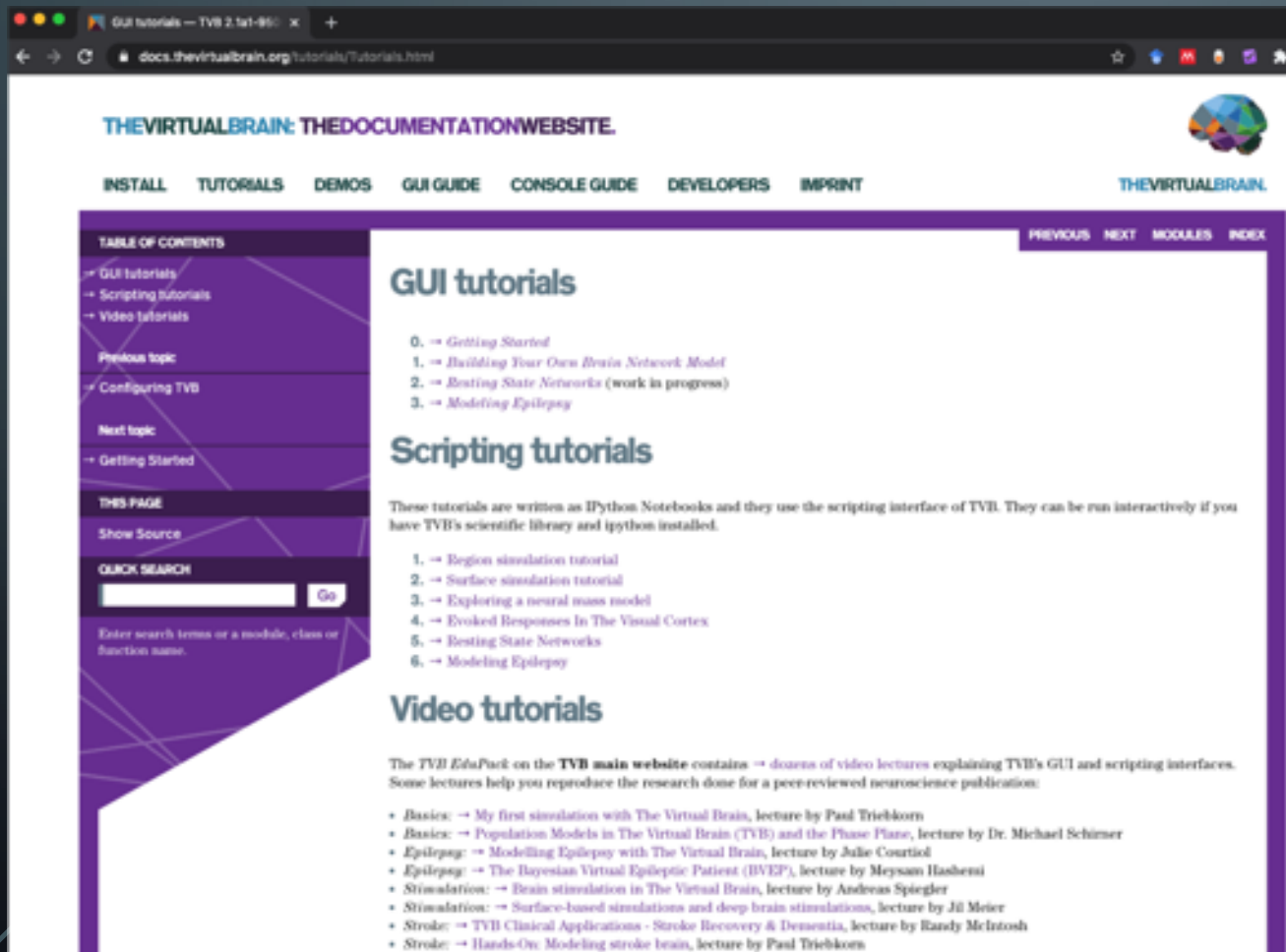


FAIR PRACTICES & THEVIRTUALBRAIN

- open-source platform (source code: github.com/the-virtual-brain/tvb-root)



FAIR PRACTICES & THEVIRTUALBRAIN



The screenshot shows the documentation website for The Virtual Brain. The page is titled "GUI tutorials" and is part of a larger documentation site. The navigation menu includes "INSTALL", "TUTORIALS", "DEMOS", "GUI GUIDE", "CONSOLE GUIDE", "DEVELOPERS", and "IMPRINT". The main content area is divided into three sections: "GUI tutorials", "Scripting tutorials", and "Video tutorials".

GUI tutorials

- 0. → *Getting Started*
- 1. → *Building Your Own Brain Network Model*
- 2. → *Resting State Networks (work in progress)*
- 3. → *Modeling Epilepsy*

Scripting tutorials

These tutorials are written as IPython Notebooks and they use the scripting interface of TVB. They can be run interactively if you have TVB's scientific library and ipython installed.

- 1. → *Region simulation tutorial*
- 2. → *Surface simulation tutorial*
- 3. → *Exploring a neural mass model*
- 4. → *Evoked Responses In The Visual Cortex*
- 5. → *Resting State Networks*
- 6. → *Modeling Epilepsy*

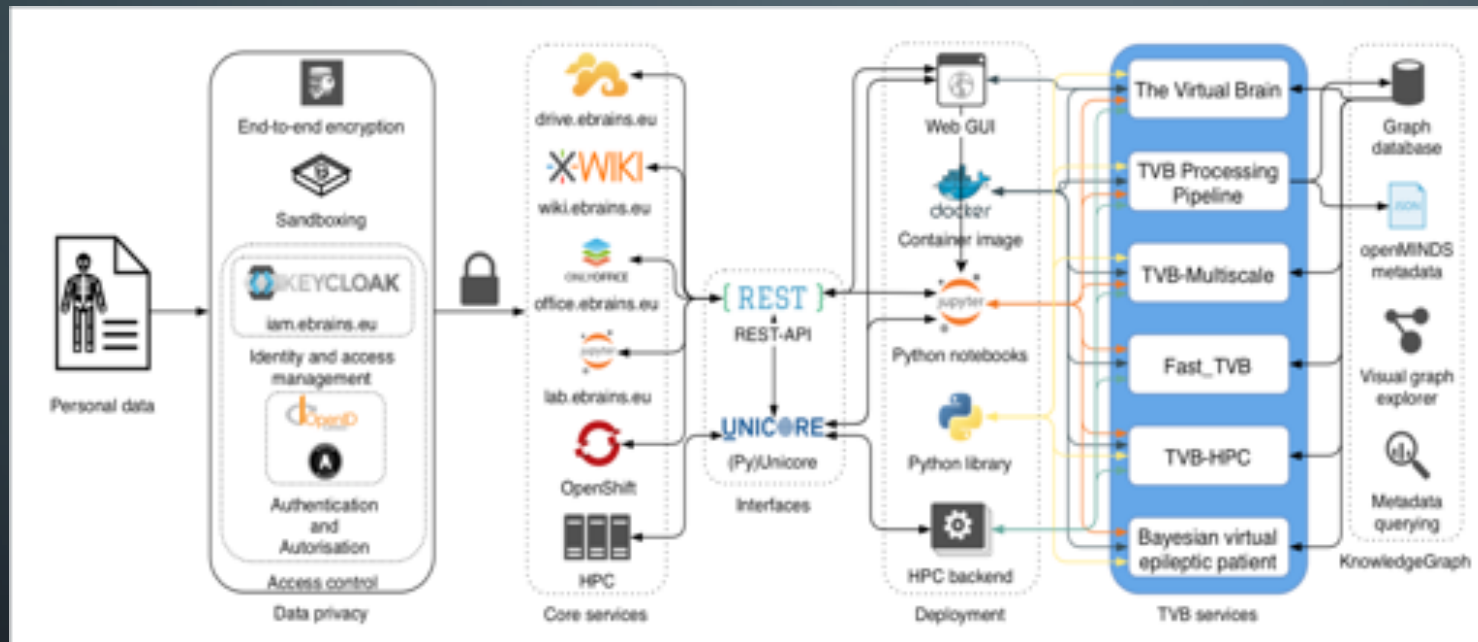
Video tutorials

The *TVB EduPack* on the **TVB main website** contains → dozens of video lectures explaining TVB's GUI and scripting interfaces. Some lectures help you reproduce the research done for a peer-reviewed neuroscience publication:

- *Basics*: → *My first simulation with The Virtual Brain*, lecture by Paul Triebkorn
- *Basics*: → *Population Models in The Virtual Brain (TVB) and the Phase Plane*, lecture by Dr. Michael Schürner
- *Epilepsy*: → *Modelling Epilepsy with The Virtual Brain*, lecture by Julie Courtiol
- *Epilepsy*: → *The Bayesian Virtual Epileptic Patient (BVEP)*, lecture by Meyssam Hashemi
- *Stimulation*: → *Brain stimulation in The Virtual Brain*, lecture by Andreas Spiegler
- *Stimulation*: → *Surface-based stimulations and deep brain stimulations*, lecture by Jil Meier
- *Stroke*: → *TVB Clinical Applications - Stroke Recovery & Dementia*, lecture by Ehsay McIntosh
- *Stroke*: → *Hands-On: Modeling stroke brain*, lecture by Paul Triebkorn

- extensive documentation, training material & support (docs.thevirtualbrain.org)

IMPROVING FAIR PRACTICES WITH TVB CLOUD & EBRAINS



Schirner et al. (2021)
arXiv:2102.05888

- collaborative & reproducible workflows
- security protocols for sensitive health data

Also see INCF Assembly 2021
Session 6: Tools & infrastructure showcase
Petra Ritter on “The Virtual Brain”

REMAINING CHALLENGES TO FAIR FOR PERSONALIZED MODELS

- Standardized model outputs, filetypes and metadata
- Which empirical data derivatives are necessary for reproducibility?
- “Standardized” processing of other neuroimaging modalities (e.g., M/EEG)
- New challenges posed by the integration of other non-brain data (e.g., genetics, metabolomics)

THE TVB TEAM



Rotman Research Institute

- Randy McIntosh
- Tanya Brown
- Noah Frazier-Logue
- Zheng Wang
- Devin Sodums
- Justin Wang

Univ Aix-Marseille

- Viktor Jirsa
- Marmaduke Woodman

Codebox

- Jochen Mersmann
- Lia Domide

Charité Berlin

- Petra Ritter
- Michael Schirner

