

Matthew Brett is the project PI; he is currently a lecturer in data science at the University of Birmingham (UoB). He trained and practiced as a doctor before moving into research; he has published in cognitive neuroscience, brain imaging methods, statistics and computing (current h-index 36). He is also an experienced scientific software developer who has significant contributions to the primary Python packages, is a Scipy core developer, and a co-author on the recent Numpy and Scipy papers in Nature / Nature Methods. Aston, Turkheimer (below) and Brett wrote the first reproducible paper in brain imaging, in 2006. Brett was a teacher-researcher in Berkeley from 2008-2017, during which time he developed new code-to-learn teaching methods for data science and brain imaging, and co-led a radical and successful undergraduate course on reproducible data science (see ref in case for support). He has much experience of leading and working in open-source communities; he co-founded the successful Neuroimaging in Python initiative (<https://nipy.org>) and led various successful software projects, including the Nibabel, a widely-used base library for brain imaging in Python. He teaches undergraduate and master's level data science using R and Python, has run his own JupyterHub cloud instance at Birmingham, and publishes his current Python course online with a permissive license. He commits to 40% of his time and will lead main project meetings, collate, develop and share teaching materials for the main FoDS / RDS courses, and coordinate development work for Open edX and JupyterHub / notebook integration.

Juan Kloppe is a specialist surgeon, educator, and data scientist. He is a leading figure in the development of modern data science material, using platforms including Coursera, Udemy, RPub, GitHub and YouTube. His Coursera course on biostatistics for healthcare professionals has nearly 100,000 learners and has been listed in the Top 50 courses of all time on the indexing platform Class Central since 2017. He is also the only lecturer with three courses on the platform; his course on computational thinking for statistical analysis using contains an honours module on deep neural networks for use in clinical healthcare and radiological diagnosis. He was the 2014 Open Education Consortium Educator of the Year. He has more than 1,400 lectures on YouTube nearing 1,500,000 views, subjects include linear algebra, calculus, graph theory, statistics, machine learning and deep neural networks, and clinical medicine. He teaches using R, Python, Julia and Mathematica; his code is freely available on Github. Juan is the Faculty representative the Senate Committee on Online Education of the University of Cape Town, and has been the Head of Research, the Head of Surgical Education, and the Head of the Acute Care Surgical Unit at Groote Schuur Hospital. He commits to 50% of his time, and will use his extensive experience with online data science learning to contribute to development and teaching of the core courses and will lead the connector course on probability.

Peter Hansen (20% time) is a senior lecturer at UoB, and specializes in computing, data science and neuroimaging. His training and early research were in physics, with a heavy emphasis on computing, after which he ran a successful consultancy company for programming and algorithms. Later he rejoined academia to work on brain imaging in Oxford, and UoB, including recent work with big data in brain imaging databanks. He has been the departmental lead for IT and a committee member for the University IT strategic board. He currently runs the UoB post-graduate (PG) brain imaging course, and has led the PG programming course. He will help build the contribution frameworks, contribute to coordination on computing development work, and teach on, contribute to the foundation and reproducible data science courses.

Valentin Danchev (20% time) is a lecturer in Computational Social Science the University of Essex. He uses large datasets and methods from computational social science to study

issues including transparency and reproducibility of scientific research, human mobility, global migration, inequality, digital-health interventions, and data sharing and reproducibility of research claims in biomedical sciences. He publishes data and code in Python and R as Jupyter Notebooks on OSF and Github, and teaches introductory computational social science and data analysis to students without prior coding experience. His courses cover basic data handling as well as machine learning, causal inference, and network analysis, using Jupyter notebooks. He will contribute to materials and teaching for main foundations course, and will lead the development of the computational social science connector course.

Federico Turkheimer (15%) is Professor of Neuroimaging (Analysis & Statistics) at King's College London (KCL). He has training in electronic engineering, Nuclear Medicine, and biostatistics (at NIH), has worked at Cambridge and Imperial College and consults on quantitative methods for industry. He studies models of brain function using statistical methods of imaging data (H-index 68). He has been teaching statistics for >20 years and has written a book "Statistics without equations." His current focus is the >1000 student KCL neuroscience distance learning programme. He will work on teaching for the core modules, and lead the multivariate methods connector course.

Brenda Williams (10%) is deputy head of Department of Education for Basic and Clinical Neuroscience, at the institute of Psychology, Psychology and Neuroscience (IoPPN), and the Faculty Assessment and Feedback lead. She pioneered the introduction of distance learning at the IoPPN and is now Programme Lead for two highly successful online MSc programmes: Applied Neuroscience and the Psychology and Neuroscience of Mental Health. She has extensive experience in online and distance learning, and is passionate about inclusive education and student engagement. She will work all courses to review and optimize materials, and will lead development of the online learning experience.

Robert Leech (10%) is Professor of Neuroimaging Analytics at KCL. His research has focused on developing and applying neuroimaging methodologies to understand the relationship between brain and behaviour. His work is inherently multi-disciplinary, with collaborations across disciplines including computer scientists, machine learning experts, cognitive psychologists, clinicians, bioengineers and physicists. He has used many programming languages (including Python and R) for data analysis. Has developed automated machine learning approaches for data exploration/inference and prediction as well as theoretical work demonstrating how these approaches can improve the reproducibility of scientific findings. He was also an editor for the journal Neuroimage's special section on reproducible neuroimaging. His teaching has been heavily inspired by the coding-to-learn philosophy, teaching coding and data analysis to cognitive psychology and neuroscience PG students for >10 years. He is module lead for neuroimaging for the applied biostatistics MSc and will be the module lead for computational neuroscience and machine learning in the new BSc in Psychology and Neuroscience. He will contribute material and teach on the core courses, and on the multivariate analysis connector course. He will also work on development of the hosting, including customization of the Open edX machinery.