

Mozilla Science Mini Grant Application

View on Screendoor at <https://mozilla-science-lab.forms.fm/mozilla-science-mini-grant-application/responses/WMOth6Ru2SEW>

Name

No response

Email

No response

Legal name

Isla Staden

Email

istaden@turing.ac.uk

Phone number with country code

+447528472170

What is your physical address (including postal code and country)?

Alan Turing Institute
British Library
96 Euston Road
London NW1 2DB

What institution are you based at currently, if any?

Alan Turing Institute

Have you previously received a grant from Mozilla Foundation? If so, list the amount and purpose of the grant.

No

Proposal Title

Brain Networks in Python

Provide a brief description of your project or event.

The Brain Networks in Python project provides open source, documented, tested and modular code to investigate the brain as a network. These analyses are reproducible, transparent and accessible to new and expert brain imaging researchers and network neuroscientists. Openly developed, this project welcomes users and contributors from all disciplines.

Total proposal budget in US Dollars (USD).

\$4960.00

How much money in USD are you requesting from the Mozilla Foundation?

\$4960.00

List your partners, if any, and the role/s they will play.

I will be mentored by Dr Kirstie Whitaker. Kirstie is a research fellow at the Turing Institute and outgoing Mozilla Fellow for Science. She is a neuroscientist who studies brain development, using graph theory to understand the brain as a network. Dr Whitaker is a two time mentor for the Mozilla Open Leadership Training Series and she has extensive skills in nurturing and supporting open projects. The Brain Networks in Python is built within the Whitaker Lab GitHub repository where Dr Whitaker is an active contributor. She will review all code and documentation and provide guidance to ensure the Brain Networks In Python project is accessible and usable for all.

I will also work with the research software engineering team at the Alan Turing Institute. The team is lead by Dr James Heatherington who has worked for 5 years as the head of Research Software Development at University College London. In his words: "Well written, readable software, written for humans to read as well as computers to execute, forms an important part of research communications and can deliver significant research impact." (<http://www.ucl.ac.uk/research-it-services/people/james>)

On which Internet health issue(s) does your proposal focus? (Select all that apply.)

- ☒ Open Innovation
 - Digital Inclusion
 - Decentralization
 - Privacy and Security
 - Web Literacy

If you are NOT a U.S. resident or citizen, will the proposed project or event involve travel to, or activities within, the United States? (This question is being asked solely for tax purposes and is not used for determining proposal acceptance.)

- ☐ Yes
- ☒ No
 - Not applicable

Are you requesting support for an event (e.g. workshop, training, hack-a-thon, or other event) or a project (e.g. open software, curriculum, or other project)?

- ☐ Event
- ☒ Other Project

Describe the issue/problem you are trying to address.

The brain made up of many regions that communicate millions of messages every second. Traditional neuroimaging analyses have considered each region separately and therefore failed to capture many aspects of brain structure and function. Graph theory allows us to analyse human brain images as a network (Bullmore & Sporns, 2009). The most popular toolbox to do brain network analyses is written in Matlab (Brain Connectivity Toolbox; <https://sites.google.com/site/bctnet>) which requires an expensive (and therefore inaccessible) license. We want to provide free and openly available software to allow all neuroscience researchers to conduct reproducible and biologically meaningful investigations of the human brain.

List key project activities (what will you do), outputs (what will be produced through your activities, e.g. products, publications, number of people trained) and outcomes (impact of your project on your beneficiaries during the grant period).

The three goals of this mini-grant proposal are to 1) unlock the specialised knowledge required to conduct graph theoretic analyses on brain imaging data by providing clear documentation, 2) make an existing code base more modular and therefore flexible enough to be used for a wider range of applications, and 3) to promote a free and open source community driven project that facilitates the highest level of scientific research. The proposed activities, outputs and outcomes are explained in detail below.

Activities

The Brain Networks in Python code was implemented in the Neuroscience in Psychiatry Network publication "Adolescence is associated with genomically patterned consolidation of the hubs of the human brain connectome" published in PNAS in 2016 (Whitaker*, Vertes* et al, 2016). The code conducts structural covariance network analyses on cortical thickness measurements from two independent cohorts and creates all figures and results tables for the published manuscript. The code and data to support the manuscript can be found at https://github.com/KirstieJane/NSPN_WhitakerVertes_PNAS2016.

As currently written, the Brain Networks in Python code is monolithic. It is very difficult to repurpose for new analyses or for different datasets. The first goal of this project will be to *refactor the code into modules*. Specifically, I will break down the conceptual steps for a network analysis and allow the user to re-combine them as they see fit. For example, one user may want to use the code only for the graph theory measures while another may use the visualisation module alone. It will be possible to link the modules together to complete an entire neuroimaging analysis from start to finish, but this workflow will not be necessary. Modularising the software will permit a more diverse user base and a much more interoperable set of code.

The second best time to add in tests for software is when refactoring. (The optimum time would have been when the code was written in the first place). As part of this 2 month project we will *add in high level and unit tests* for the network analysis code. Testing will give users confidence in our code and allow us to implement continuous integration. We are already developing in a version controlled environment on GitHub, but as the project develops, we will always be able to verify the accuracy of the modules.

We will also add *new functionality* to the code. There is currently no way (in this codebase) to investigate the difference in brain networks between two groups. For example, a key question in the field of biological psychiatry is whether the brains of people who have a diagnosis of depression are differently connected to the brains of healthy control participants. We will add a module that permutes the original data and builds up a null distribution so that differences between the two real groups, for example people with a diagnosis of depression and healthy controls, can be statistically tested.

Along with integrated tests, *documentation* will allow the Brain Networks in Python code to be used by many people. We will provide example data and clear tutorials to make it easy for users to install and get up and running with the code. These tutorials will be in the form of Jupyter notebooks that can be either downloaded and run locally or run in a MyBinder (<http://mybinder.org>) instance online. An example Jupyter notebook that is currently in development can be found at https://github.com/WhitakerLab/BrainNetworksInPython/blob/master/Example_JupyterNotebook.ipynb.

Finally, we will *release our code as a Python package* that can be easily installed via pip (<https://packaging.python.org>) or conda (<https://conda.io>). We will ensure that all released versions of our code have digital object identifiers from Zenodo (<https://zenodo.org>).

Outputs

- * MIT licensed, documented and tested code that is easy to use and re-use.
- * Modular code that can be repurposed for individual use cases.
- * An easy to install python package that can be cited in future publications.
- * Version controlled code that permits rigorous reproducible neuroscience.
- * An open and inclusive community dedicated to supporting new and veteran users and contributors.

Once a community grows around the project we hope to inspire additional integrations. For example collaborating with Mozilla Global Sprint project Cytoscape js (<http://js.cytoscape.org>) to provide interactive visualisations of the networks. We could also harness the power of Docker containers promoted by BIDS (Brain Imaging Data Structure) Apps (<http://bids-apps.neuroimaging.io>) to make it even easier for analyses to be reproducible and accessible. The code currently provides very standard graph theoretic measures, but we would hope to integrate novel computations such as the network versatility (Shinn et al, 2017; <https://github.com/mwshinn/versatility>).

Outcomes

The Brain Networks in Python project will make it easier to conduct reproducible analyses of neuroimaging data. We will focus on smoother onboarding for new users (for example graduate students and postdoctoral researchers) to conduct network analyses. We will ensure that early career researchers from around the world can conduct the most rigorous and state of the art analyses of the human brain.

The outcomes of this project during the grant period will be to increase the number of researchers who have the confidence to use and contribute to free and open source software. By supporting our community of early career researchers we will build towards a critical mass of scientists who require open, reproducible and accessible research as a standard for the future of research.

Provide key indicators you plan to use to measure project outcomes and source of data.

* Number of contributors. We will track the number of commits in any branch of the project along with any comments on issues or pull request reviews as contributions using the Let All Build a Hatrack tool (<https://labhr.github.io>). We will acknowledge these contributors publicly within the GitHub project repository using Kent Dodd's All Contributors emoji key (<https://github.com/kentcdodds/all-contributors>).

* Package downloads. We will track the number of downloads of our Python package using the <http://pypi-ranking.info> website.

* Scientific citations of the code. We will track the Zenodo DOI for each version of the code to see how often the Brain Networks in Python code is used in academic publications. There are three projects that Kirstie Whitaker is already involved in that will cite the code on submission for publication by the end of 2017.

Explain who will benefit from the project.

Early career researchers will find friendly support for their academic work.

Complex network analysts who integrate their models will experience wider adoption of their work.

Reproducible and version-controlled software ensures efficient knowledge transfer. Senior scientists as expertise will not leave their group when one member advances to a new position.

Describe the geographic location(s) in which this project will be implemented.

London, UK

List any risks or challenges that may affect the overall success of your project, and note how Mozilla and/or others can help you to overcome these challenges.

Although the code has been used very successfully by Kirstie Whitaker in previous publications, it is difficult to understand for external collaborators. I have been working hard to implement the code but it has been necessary to walk over to Kirstie's desk and ask her to explain (for example) naming structures and project organisation.

It is easy to see how this situation developed. The person who originally wrote the code understands it so well that she has expert blindness to the challenges that fresh eyes may experience. It is imperative that we have user input from the beginning. The Mozilla Science Lab network will provide excellent opportunities to report on project development and receive feedback from the community.

One of the greatest challenges for the project is that of finding time to build documentation, add in a test suite and refactor the code to be more modular and therefore more widely applicable. This mini-grant will be particularly valuable in levelling up the code base to make it easy to use for international standard publications in the field of network neuroscience.

Is this a new project or a continuation? If new, please describe your qualifications to initiate the activity. If continued, please describe your accomplishments to date. Feel free to include links to articles and documents online which highlight your recent work.

The project is a continuation of the work that Dr Kirstie Whitaker published in 2016 (Whitaker*, Vertes* et al, 2016). The code is openly available on GitHub (https://github.com/KirstieJane/NSPN_WhitakerVertes_PNAS2016). For this work, Dr Whitaker and her collaborator Petra Vertes were named as 2016 Global Thinkers by Foreign Policy magazine.

Dr Whitaker's CV is available at our lab GitHub repository:

https://github.com/WhitakerLab/WhitakerLabProjectManagement/blob/master/Kirstie-Whitaker/Whitaker_CV.pdf

I am a research assistant in the Whitaker Lab at the Alan Turing Institute. I have a BSc in Mathematics from the University of Cambridge which included specialised training in graph theory and topological analysis.

I have advanced coding skills in Matlab and Python, including expertise in the matplotlib, networkx, numpy and pandas libraries that are used by the Brain Networks in Python project. I am a proficient git & GitHub user, and have experience writing academic documents in LaTeX.

I am already working on the project described in this proposal and a list of my code contributions can be found at <https://github.com/WhitakerLab/BrainNetworksInPython/commits?author=Islast>.

Mozilla works in the open. How will you document and share your project progress with the community.

The Brain Networks in Python project is hosted on GitHub: <https://github.com/WhitakerLab/BrainNetworksInPython>.

We follow the Mozilla open leadership training best practice principles (<https://mozilla.github.io/open-leadership-training-series>). The Brain Networks in Python project has a README file, a code of conduct and a guide for contributors. We manage any updates through issues and transparently review all pull requests.

Within the Whitaker Lab we work as openly as possible. All the weekly meetings I have with Dr Whitaker are available (<https://github.com/WhitakerLab/WhitakerLabProjectManagement/blob/master/Isla-Staden/Weekly-Meetings.md>). We are dedicated to sharing work in progress so that others can provide help and feedback. In my first week as a member of the Whitaker Lab I contributed to the Whitaker Lab blog (<https://whitakerlab.github.io/blog/Hello-Earthling/>) and I am greatly looking forward to the opportunity to engage further with the brain networks community through social media and further blog posts.

How will you continue work on this project beyond this funding period?

Dr Whitaker will continue to lead the Brain Networks in Python project during her five year fellowship at the Turing Institute. I hope to find funding to continue my contributions beyond the end of the grant period possibly through the Alan Turing Institute's seed funding program. A requirement for seed funding is that a product or publication can be completed in six months. The Mozilla mini-grant would give me time to build community around the Brain Networks in Python software and therefore build a successful application to the internal Turing funding body.

Importantly however, I may not be needed to continue on the project. If our two month summer program of refactoring, modularisation, testing, documentation and community building goes to plan we will have a powerful open source community that can continue to develop without my input. I hope to remain involved in the project, but I also hope that I am not necessary. A key deliverable of this mini-grant would be a Brain Networks in Python package that can sustain itself in a stable form.

Please upload an itemized project budget using the template linked below the button.

- StadenMiniGrantBudgetTemplate.xlsx (49.2 KB)
<https://screendoor.dobt.co/attachments/ibfi9uCkzuG-eZ2n3hFBVVTvmzh0C26k/download>

How did you hear about the mini-grant opportunity?

- ✓ Twitter
- ✓ Blog
- Friend

Do you want to be added to our Mozilla Science Lab mailing list?

- ✓ Yes, please!
- Not at this time.