Phase II Competition - Finalist Application Form

Created: 11/11/2016
Last updated: 11/22/2016

In this form, the six finalist teams for the Open Science Prize are asked to describe the work they have done to develop their prototypes, and make the case for why they should be considered for the phase II Prize.

The information you submit on this form will be considered alongside the prototype you have developed in deciding which team will receive the Phase II Prize.

Please note that, unless otherwise stated, the information you submit on this form will be made available publicly via the Open Science Prize website (under a CC BY 4.0 license), so that it can be assessed as part of the public voting process. All fields, except the final box for additional information are mandatory.

Your application must be completed by 11:59pm Pacific Standard Time on 21 November 2016 . You may edit this form as many times as you like before the deadline.

If you have any questions about the Prize or the review process or if you would like to provide any further information that you would not wish to be made public, please contact David Carr (d.carr@wellcome.ac.uk) or Elizabeth Kittrie (elizabeth.kittrie@nih.gov). Any technical questions regarding this form or the web platform should be directed to (openscience@wellcome.ac.uk).

Page 1

Executive Summary

Please provide a brief Executive Summary of no more than 150 words for the public voting page on the Open Science Prize website. This should be suitable for an informed lay audience, and should briefly describe your prototype and why it should be considered for the phase II Prize. [150 words]

Mental and neurological disorders pose major medical and socioeconomic challenges for society. Understanding human brain function and disease is arguably the biggest challenge in neuroscience. To address this challenge, smaller but sufficiently complex brains like that of the fruit fly have been increasingly used for investigating the mechanisms of human neurological and psychiatric disorders, such as Epilepsy or Parkinson’s disease, at molecular, cellular and circuit level.

The Fruit Fly Brain Observatory (FFBO) is an open source software platform that stores and processes data related to the neural circuits of the fly brain including location, morphology, connectivity and biophysical properties of every neuron; seamlessly integrates the structural and genetic data from multiple sources that can be queried, visualized and interpreted; automatically generates models of the fly brain that can be simulated efficiently using multiple Graphics Processing Units (GPUs) to help elucidate the mechanisms of human neurological disorders and identify drug targets.

Weblink for prototype

Please provide the public URL for your prototype tool or service (this will be viewed by the public for purposes of public voting):

http://fruitflybrain.org

Your Prototype

Purpose and need

Please provide a brief summary description of the purpose of the prototype you have developed and the key challenges or needs it is seeking to address [200 words]

FFBO is a unique public resource for Drosophila brain data storage, composition, simulation and visualization. FFBO supports fundamental and translational neuroscience research, turning insights into the causes and mechanisms of human neurological diseases, gained from fruit fly disease models, into novel therapies.

Currently, researchers aiming to devise computational models of neural circuits in the fly brain are faced with a fragmented array of connectome and genetic data sources. There are virtually no resources for configuring and simulating computational models of the fly brain to test neurological disease hypotheses. Integration of data from various genetic, anatomical and clinical databases into a testable computational disease model is currently performed in ad hoc fashion with questionable reproducibility and reusability.

FFBO addresses these issues first by providing a comprehensive repository for biological data and computational models of healthy and diseased neural circuits, which is tightly integrated within a powerful simulation and visualization platform.

Further, FFBO enables the investigation of mechanisms underlying neurological disorders and potential therapies by configuring, simulating and interrogating fly brain models developed by different researchers. These models can then help predict the effects of pharmaceuticals upon neural circuit functions, significantly shortening the amount of time it takes to develop new drugs.

Please summarise the work you have taken forward to develop your prototype since you were awarded the Phase I Prize in April 2016.

i. Progress

The key milestones achieved and the extent to which the goals and challenges you set out to address in your original application were delivered [400 words]

We have achieved all the goals and challenges that we set out in our original application. Our prototype achieves the following milestones towards our goals:
- We developed NeuroArch, a database for storing, querying and manipulating biological and executable circuit data via a unified API.
- We produced a data model for representation of both biological data and executable circuit data in NeuroArch.
- We provided an API that makes it easy to import and export biological circuit and executable model datasets, query data at different levels of structural granularity or abstraction, and algorithmically construct and dispatch executable models to Neurokernel, the brain emulation engine we developed.

With the development of NeuroArch, we have proceeded to complete and surpass the stated aims of the Phase I application well beyond our expectations. In addition to the original milestones we have also:

- Created NeuroNLP, a web-based portal for navigating fruit fly brain circuit data stored in NeuroArch through plain English.
- Built NeuroGFX, a graphical functional explorer, a highly intuitive set of tools for executing neural circuit models with Neurokernel.
- Defined a software architecture for flexible integration of modules that supports NeuroNLP and NeuroGFX, as well as new applications for creating new healthy and diseased fly brain circuit models.
- Developed NeuroAPPs, early iterations of executable disease models.

NeuroNLP enables in-depth exploration and investigation of brain structure, using intuitive natural language queries. This simple yet powerful interface replaces the usual, cumbersome checkboxes and dropdown menus prevalent in today's neurobiological databases, enabling easy access to structural data.

NeuroGFX provides an intuitive graphical interface to leverage the power of the computational infrastructure of NeuroArch and Neurokernel and allows visualizing execution results in the context of biological structure. This provides an environment where computational researchers can present configurable, executable neural circuits, and experimental scientists can easily explore circuit structure and function ultimately leading to biological validation.

New applications can be created for different disease models, such as what is shown under “NeuroAPPs”. To this end we have

- Created an epilepsy model which can be used to reproduce experimental observations and to verify working hypotheses.
- Developed an interactive demonstration of the poorly-understood effect of the retinal degeneration seen in Parkinson's disease.
- Created a retina app to study the potential of rescuing retinal degeneration by optogenetic means.
- Emulated pathological states of the olfactory system due to excess release of the GABA neurotransmitter, a phenomenon observed in Parkinson’s disease.

---

ii. Team Contributions

The contributions of the team members to the development of the prototype [200 words]

As part of the development process, members from each university team were able to meet in New York to coordinate and map out the long term software development goals, as well as kickstart the initial platform development, enabling more effective and cohesive development for the duration of the project.

The Columbia University team implemented and improved the NeuroArch data model and API, extended the Neurokernel platform to support the new data model and developed NeuroNLP and NeuroGFX. They worked closely with The University of Sheffield team to design and develop the modular architecture, forming the backbone of the FFBO platform, supporting distributed development of new modules. The National Tsing Hua University team provided morphological and meso-scale connectomic data, expanding the utility of the platform as a whole. Importantly, with close collaboration, each team has been able to contribute according to their expertise, while sharing knowledge in skills in each domain. This culminated with each group building their own disease models, such as with retinal degeneration, Parkinson's disease and Epilepsy.

---

iii. Significant Achievements

Any significant achievements or key success metrics you wish to highlight - this might include, for example, numbers of users, key endorsements or engagements with users, new partnerships, external funding, and so forth [200 words]

The FFBO is the world's first open science platform to integrate, within a single database, genetic, anatomical, neurophysiological data with computational models, all of which are coupled with powerful executable circuit generation, numerical simulation and user-friendly query and visualization tools.

FFBO features NeuroNLP, a unique natural language interface for querying the fly brain database, which already supports three languages. A user-friendly graphical interface provides powerful interactive data integration and visualization functionality of neural circuits. A unique feature and major achievement of FFBO is the programming interface NeuroGFX that automatically translates biological and modeling data into executable code that can be run on a local or cloud-based GPU servers. The FFBO is publicly available and accessible from any modern web browsers, including those running on smartphones.

The core supporting technologies of the platform were demonstrated earlier this year at our first ever Columbia Fruit Fly Brain Hackathon. Presentations are planned for the 2017 Columbia Workshop for Brain Circuit, Memory and Computation.

The development of the FFBO platform is partly supported by grants from BBSRC (UK), NSF (USA) and Ministry of Science and Technology (Taiwan), and by time contributions of researchers in labs at Cambridge, Oxford, Rockefeller, Stanford and Washington University.

Learning Points

Please briefly highlight any key learning points you took from the work that you undertook to develop your prototype [200 words max]

Development of an open science platform has to be a collaborative effort involving multidisciplinary teams with public/end-user engagement. Our current progress in delivering the FFBO prototype underlines the effectiveness of an engagement strategy which delivers technology that addresses the immediate needs of the research community. The success of the hackathon event has further highlighted the future requirement to deliver training and support for end-users. Whilst end users and developers (currently) span three continents, our NYC meetings were highly effective for refactoring of code as well as merging of cross over functionality such as the integration of the disease models. Remote working on well planned modular and scalable units of function were complemented by these infrequent face to face meetings. The project’s current success has been highly dependant on feedback from strategically selected research groups. Moving forward we will need to formalize the feedback process as we engage a wider user base. Our open software approach and use of Request for Comments (RFCs) and community tools such as Github will allow this, ensuring that we can target a wider range of user communities.

Case for Phase II Prize

Please make the case for why your prototype should be considered for the Phase II Open Science Prize against the following key criteria [100 words each]:

i. Impact

The current and potential future impact of the tool or service in terms of advancing research and generating health and societal benefit

The FFBO platform challenges existing models of knowledge creation in neuroscience. It provides the framework and tools to build and simulate fly models of neurological diseases using data and models developed by the community of researchers at large.
By providing a model architecture upon which researchers can build, share, compare, refine and validate models of neuropil compartments, constituent circuits and connectivity maps, FFBO will unify and support the research efforts of labs around the world, accelerating the pace of discovery and the translation of fundamental neuroscience research into drug, cell and gene therapies. FFBO is also a unique educational tool.

ii. Innovation

The degree of innovation associated with the tool or service

Executable brain models will be key to elucidating the function of brain circuits in health and disease. Our platform features an integrated graph database that represents both biological information and executable brain circuit models. It can integrate independently developed neural circuits and provides highly accessible natural language and interactive graphical interfaces that are tangible to both neurobiologists and computational neuroscientists alike. The FFBO prototype pioneers the key steps towards creating a paradigm of scalable computer aided design (CAD) tools for neural circuits; it provides pivotal methodologies and technologies for successfully bridging the gap between brain structure and function.

iii. Utility

The level of demand and utility associated with the proposed service or tool

Our platform has great utility for the fruit fly research community. Importantly, the technologies we have developed are readily scalable to requirements for studying brain function in higher organisms. This will have a ripple effect on neuroscience communities outside the insect world, including but not limited to zebrafish, mice and primates. Furthermore, our platform is central to neurobiologists, computational neuroscientists and neuromorphic engineers as it enables collaborations across these disciplines. With its seamless user accessibility, the platform lowers the barrier of entry into the field and provides a great educational and research resource.

iv. Feasibility & Technical Merit

The feasibility and technical merit of the prototype

Our platform leverages leading-edge technologies, including massively parallel computation, integrated graph databases, standard API based scalable system architecture for modeling of a whole brain and use of natural language processing for querying biological and modeling data. The latter capability stands in stark contrast to cluttered menus and submenus prevalent in other contemporary data portals.

We have broken down the technical barrier of tightly integrating biological and computational modeling information, and are opening up new avenues to explore neural functions, far from the individual focus of small labs and closed development environments.

Development & sustainability plan

Please briefly describe your vision, and any tangible steps you have taken, to develop your prototype into a sustainable tool or service that advances the goals of open science [400 words]

We recognize that fostering open and active community interaction is the key to a impactful project. Going forward we will continue in the vein of opening up the FFBO to developers and scientists alike.

Building on the success of our first Fruit Fly Brain Hackathon in March 2016, we will be holding further hackathons, to introduce developers to the latest FFBO achievements, and encouraging development of new apps that exploit and improve our current and powerful APIs and our various data sources. To receive user feedback from the neurobiology community, we will continue organizing Columbia BCMC workshops with the goal of bringing together researchers interested in executable models neural computation/processing of the brain of model organisms.

We have focused on improving the accessibility of the software by providing pre-built machine images that can be easily installed and run on cloud computing services, allowing anyone to get involved regardless of restrictions imposed by lack of locally available hardware resources.

We will continue to encourage the development of third party applications, by developing and releasing the code on online repositories, providing ample documentation and support. The software has been designed from the bottom up to be modular and expandable, supporting both publicly and privately hosted databases, opening up the possibilities of community integrated ‘NeuroAPPs’.

To finalize the development pathway, we will be publishing RFCs along with documented demonstration applications. RFCs encourage community involvements to shape the progress and feature requests of the platform.

We have envisioned two models of hosting the service. For the first model, FFBO provides docker images of the FFBO software base with instructions on how it can be installed in house or in the cloud. In either case, the user is responsible for the installation and updating all of the code. FFBO plays the role of a clearinghouse for software development and distribution. The second model calls for providing a service for all interested users via a web browser. This model has a very low barrier of entry but requires maintaining and providing at cost computing resources for code execution for all interested parties. Such a service, while desirable requires substantial additional funding. The advantage of the latter model is that there would be an open central repository for all information about the fruit fly brain including executable brain circuits.

Final comments

Please use the box below to provide any further information you would like to add, that has not been addressed in the questions above [200 words]

We believe that the time is right to bring the fruit fly brain community together with the goal of elucidating its function in health and disease in the next 10 years. A major international activity as part of an International Brain Station might be a good starting point. We are not asking here for additional funding support but rather for a co-ordination of the funding organizations (NIH, NSF, Wellcome Trust, HHMI, etc.) that support fruit fly brain research and ask for a mandate that all data and code generated from this research are published in an easily accessible form from a single integrated data repository.