Research Ideas and Outcomes 2: e9113
doi: 10.3897/rio.2.e9113

Small Grant Proposal

Open Neuroimaging Laboratory

Katja Heuer[‡], Satrajit S. Ghosh[§], Amy Robinson Sterling[|], Roberto Toro[¶]

‡ Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
§ MIT, Cambridge, United States of America
| EyeWire, Cambridge, United States of America
¶ Institut Pasteur, Paris, France

Corresponding author: Roberto Toro (rto@pasteur.fr)

Reviewable v1

Received: 06 May 2016 | Published: 08 May 2016

Citation: Heuer K, Ghosh S, Robinson Sterling A, Toro R (2016) Open Neuroimaging Laboratory. Research Ideas and Outcomes 2: e9113. doi: 10.3897/rio.2.e9113

Abstract

There is a massive volume of brain imaging data available on the internet, for different types of persons and animals, capturing different types of information such as brain anatomy, connectivity and function. This data represents an incredible effort of funding, data collection, processing, and the goodwill of thousands of participants. Its analysis should allow generating insight on critical societal challenges such as mental disorders but also on the structure of our cognition.

Sharing this data is great, but it's not enough: to transform data into insight, we need access to a huge neuroinformatic infrastructure that would allow us to query it, store it, pre-process it and analyse it. But it's not only computers and code, you also need people. Neuroimaging data requires a substantial amount of human curation, visual quality assessment and manual editing. All too often, the inability of small research groups, individual researchers or citizen scientists to meet these requirements will prevent/ discourage them to explore the data or will force them to discard large amounts of data because of an impossibility to embark into lengthy manual interventions.

We intend to build the framework for an Open Neuroimaging Laboratory that will enable distributed collaboration around annotation, discovery and analysis of publicly available brain imaging data. We will build this framework using browser based applications allowing individuals to work together without having to download any data or install any software. By working together in a distributed and collaborative way, sharing our work and our analyses we should improve transparency, statistical power and reproducibility. Our aim is to provide to everyone the means to share effort, learn from each other, and improve quality and trust in scientific output. The framework we propose will leverage and extend the spirit of the EyeWire game to a distributed, open laboratory for collaboration in brain imaging.

Keywords

Neuroimaging, Collaboration, Open Science, Co-edition, Open Science Prize, Magnetic Resonance Imaging, Segmentation

Introduction

Data sharing is great. There is a massive volume of magnetic resonance brain imaging data available on the internet, for different types of persons, different types of animals, capturing different types of information such as brain anatomy, connectivity and function. This data represent an incredible effort of funding, data collection, processing, and the goodwill of thousands of participants. The analysis of this data should allow generating insight into critical societal challenges such as mental disorders but also into the structure of our cognition, into ourselves.

However, data sharing is not enough. It's not enough to be able to **download** data: if you want to **transform** data into insight, you still need access to a huge neuroinformatic infrastructure that would allow you to query it, store it, pre-process it and analyse it. And it's not only computers and code, you also need people. Neuroimaging data processing requires a substantial amount of human curation, visual quality assessment and manual editing. All too often, the inability of small research groups, individual researchers or citizen scientists to meet these requirements will discourage them to explore the data or will force them to discard large amounts of data because of an impossibility to embark into lengthy manual interventions. Sharing data is not enough, you need a big lab to be able to put data to good use. However, instead of curating shared data within each lab, it would be of greater benefit to the community to perform such curation in a collaborative setting. This would reduce redundancy in effort, increase cooperation, and help generate consensus. And even in cases when there is no consensus, collaborative curation and editing would generate challenges which the domain needs to address. By operating in silos, all these benefits are lost.

Our aim is to make possible an **Open Neuroimaging Laboratory** of the size of the planet, minimising the computational requirements for accessing the wealth of shared data to a web browser. By working together in a distributed and collaborative way, sharing our work and our analyses, we should improve transparency, statistical power and reproducibility. Our aim is to provide to everyone the means to share effort, learn from each other, and improve quality of and trust in scientific output.

We propose to develop a web framework that will illustrate and demonstrate the feasibility of our project. Our framework should facilitate:

1. **Data discovery:** We will provide tools to allow users to create shared catalogues of online brain MRI data. Initially, the catalogue will index data from online projects such as Oasis brains, ABIDE, Human Connectome Project, OpenfMRI, COBRE, Brain Catalogue, or brain MRI data available in Figshare, Zenodo, GitHub, and scientific journals. But finally, everybody can index data from their labs.

2. **Data annotation:** Data will not only be indexed, we will provide a graphical interface to allow users to collaborate online in the analysis of and discussions around this data. To illustrate the approach, we will develop a web application – **BrainBox** – to allow sandboxed collaboration. As a first example of such interaction, we will focus on manual segmentation of brain structures. A functional prototype is available at http://brainbox.pasteur.fr.

3. **Database-wise analyses:** The collaborative editing of brain MRI data will allow us to perform – online – analyses across many different datasets. By putting together all the data available online we should be able to provide the community with excellent statistical power to test new hypotheses and verify the reproducibility of results in the brain MRI literature.

In the longer run, such a framework will allow collaboration across different tools that require social communication and interaction. In the future, artificial intelligence agents or platforms can then learn from such interactions to enable better automation of such processes.

BrainBox

To illustrate the principle of our long term project, we propose to develop BrainBox: a web application for the manual segmentation of any structural brain MRI dataset available online. We have created a first proof-of-concept of our collaborative segmentation tool which you can try:

1. Go to http://brainbox.pasteur.fr and enter the URL of any structural MRI available on the internet. You can try, for example, this Human magnetic resonance imaging dataset from FigShare http://files.figshare.com/2284784/MRI_n4.nii.gz, or this Lion MRI from Zenodo https://zenodo.org/record/44855/files/MRI-n4.nii.gz. For the moment, BrainBox can only read the nii.gz format (one of the most widely used neuroimaging file formats). In the version we propose to build, the datasets will be added to your user's profile where you will be able to tag them (age, sex, species, disease status, project, etc.) and organise them into categories and projects,

2. BrainBox provides a basic online viewer which you can use to explore any online dataset. It also provides a set of basic tools which allow you to manually segment

and annotate different brain regions as well as a simple chat. These annotations are automatically saved, and will be there the next time you connect,

3. Now, invite one or more friends to connect to the same dataset you are using. The segmentation tool provides real time interaction: different users using BrainBox at the same time will be able to see each other's work, and discuss using the chat.

In the version we propose to build, we plan to implement an associated viewer which will produce 3D meshes of the structures you created. We also want to make it easy to submit segmentations to Zenodo to get a permanent reference through a DOI.

BrainBox will provide the means to create a layer of collaborative annotation over all the available MRI data without having to rely on a centralised data repository. We will not make copies of the original datasets, but only store the users' segmentations, annotations and the URL of the original dataset. BrainBox could then be used to segment any dataset that a user is able to access: if you have the rights to download an MRI dataset to your computer, then you can also edit it with BrainBox.

Our aim for the version of BrainBox that we will release by the end of the project is to improve the user experience, ensure its usability in multiple platforms and include additional MRI data formats to expand the universe of data we can access. We will pack BrainBox as a widget, so that it can easily be added to existing web pages. By adding BrainBox to your web page, you will help build a large, distributed network of brain imaging researchers from academia, citizen scientists or patients, fostering open data sharing and collaboration.

Advancement of Open Science

Our project will help transform the massive amount of static brain MRI data readily available online into living matter for collaborative analysis. We will allow a larger number of researchers to have access to this data by lowering the barriers that prevent their analysis: no data will have to be downloaded or stored, no software will have to be installed, and it will be possible to recruit a large, distributed, group of collaborators online.

Impact

Even large neuroimaging laboratories are unable today to invest into the manual editing and curation of data. We will provide a tool to facilitate and encourage the creation of distributed teams of researchers to collaborate together in the analysis of open data. In addition to the practical output of this work (i.e., the human curation of large amounts of open MRI data), our project should plant the seed for the growth of a community based on collaboration instead of competition. Further, aggregating annotated brain imaging data resources provides an improved way of searching for information and greater potential for reuse.

Innovation and Originality

Whereas many online platforms exist for co-editing text (Wikipedia, Google docs, Hackpad, etc.), the platform we propose will be to the best of our knowledge the first example of a co-editing platform for brain imaging. We foresee that such a platform can provide significant benefits: the amount of curated data available online can increase exponentially, there is greater trust in the data due to collaborative efforts, and it enables greater reproducibility by reducing redundancies and inefficiencies in processing imaging data. While a utopic platform will do away with any human curation or editing, we don't see this happening in the near term. As the amount of publicly available data grows, platforms such as BrainBox will allow increased collaboration, more effective hypothesis testing, and improved reproducibility.

Technological Viability and Resource Feasibility

The web has become the largest platform for collaboration today, and standards such as WebSockets and WebRTC have enormously facilitated the possibility of this collaboration to become real time. Our project builds on top of these standards, which ensures its technological viability. The resources required for this project – the MRI data – are already available online, but in a static form that makes them hard to analyse. Our project changes that. To demonstrate feasibility, we have developed a prototype for collaborative segmentation of brain MRI. In this prototype the data viewer (using the standard 3 orthogonal slice views) was coded in javascript and jQuery, and the layer of real-time interaction was built using WebSockets with Node.js in the server side. The annotations and segmentations are stored in a MySQL database. Many additional elements are still needed including connection to data resources, improvement of collaborative elements such as a better chat, but also discussions and offline modes of use.

The Open Science Prize Phase 1 will allow us to take this web application from prototype to version 1, improving the user experience, the stability and the documentation. Retaining our project for Phase 2 would allow us to further the type of analyses beyond the segmentation of structural MRI data and to explore the development of interfaces for database-wise statistical analyses. Our code will be released open source on GitHub to facilitate bug correction, extension and maintainability.

Bios

Satrajit Ghosh is a principal research scientist at the McGovern Institute for Brain Research at MIT and an assistant professor in the Department of Otolaryngology at Harvard Medical School. His research areas include neuroinformatics, applied machine learning, neuroimaging, and speech communication. He focuses on enhancing interoperability across brain imaging, clinical data mining, mobile health, and translational neuroimaging.

He is a lead architect of the Nipype project, an ardent proponent of distributed Web solutions for data sharing, querying, and computing, and a strong believer in solving problems together. Satrajit Ghosh has been a part of the INCF task-force on Standards for Data Sharing and has organized numerous workshops, hackathons, and coding sprints around neuroscience.

Katja Heuer is an artist, a philologist and a neuroscientist. Having studied the Fine Arts, English and French Studies in Paris and Leipzig, she is now a PhD student at the Max Planck Institute for Human Cognitive and Brain Sciences. While focusing her research on the development of the human brain and its connectome, she is also working on projects at the interface between art and neuroscience. Projects Katja Heuer worked on comprise among others The Art of MR. Staging the Connectome, Fire & Wire and Surf the connectome. 3D functional and structural brain networks in the cloud.

Amy Robinson (USA PI) is the Executive Director of EyeWire, a game to map the brain that began at MIT. EyeWire crowdsources neuroscience, challenging hundreds of thousands of players around the world to solve 3D puzzles, which actually map out neurons. This allows neuroscientists to chart synaptic connections and model circuitry. She has advised The White House OSTP and the US Senate on crowdsourcing and open innovation. Under her leadership, EyeWire's neuroscience visualisations have appeared at TED and in Times Square NYC. She also helped create the world's first neuroscience virtual reality experience, and curates the NIH 3D Print Exchange Neuroscience collection, which features several 3D printable neurons discovered by EyeWire gamers. Amy Robinson has written for Vice, the BBC, Nature, and Forbes, and writes the Neurotech series for Scientific American in partnership with MIT. She tweets @amyleerobinson, and is a long time TEDster – founder of the TEDx Music Project, a collection of the best live music from TEDx events around the world. Fast Company credits Amy Robinson with "making neuroscience into a playground for the hot tech *du jour*" and she was named to the Forbes 30 Under 30 in 2015.

Roberto Toro (EU PI) is leader of the group of theoretical and applied neuroanatomy at the Department of Neuroscience of the Institut Pasteur in Paris. He is interested in the development and evolution of the brain, which he studies through mathematical modelling, magnetic resonance imaging and genetics. He has worked on modelling the global organisation of the sulcal anatomy, created biomechanical simulations of neocortical folding, developed methods to measure cortical folding, and studied the way in which genetic and environmental factors determine neuroanatomical diversity. Roberto Toro has led the development of several web tools for online collaboration in neuroscience including BrainSpell, the Brain Catalogue, and MicroDraw.

Supplementary materials

Suppl. material 1: Animation of BrainBox's workflow

Authors: Katja Heuer, Satrajit Ghosh, Amy Robinson Sterling, Roberto Toro
Data type: Animated GIF
Brief description: Animation showing the different functionalities of BrainBox: opening a Magnetic Resonance Imaging volume, viewing it, and editing it collaboratively online.
Filename: animation.gif - Download file (3.39 MB)

Suppl. material 2: BrainBox Cartoon

Authors: Franka Barba
Data type: JPG file
Brief description: Cartoon showing the idea of distributed collaboration on Magnetic Resonance Imaging data.
Filename: cartoon.jpg - Download file (784.74 kb)