ARTICLE
Teaching with Big Data: Report from the 2016 Society for Neuroscience Teaching Workshop

William Grisham[1], Joshua C. Brumberg[2], Terri Gilbert[3], Linda Lanyon[4], Robert W. Williams[5], & Richard Olivo[6]

[1]Psychology Department, University of California, Los Angeles (UCLA), Los Angeles, CA 90095-1563; [2]Department of Psychology, The City University of New York (CUNY): The Graduate Center and Queens College, Queens, NY 11367; [3]Application Scientist, Allen Institute for Brain Science, Seattle, WA 98109; [4]Executive Director, International Neuroinformatics Coordinating Facility (INCF), Stockholm, Sweden; [5]Genetics, Genomics and Informatics, University of Tennessee, Memphis, TN 38163; [6]Biological Sciences, Smith College, Northampton, MA 01063.

As part of a series of workshops on teaching neuroscience at the Society for Neuroscience annual meetings, William Grisham and Richard Olivo organized the 2016 workshop on "Teaching Neuroscience with Big Data." This article presents a summary of that workshop.

Speakers provided overviews of open datasets that could be used in teaching undergraduate courses. These included resources that already appear in educational settings, including the Allen Brain Atlas (presented by Joshua Brumberg and Terri Gilbert), and the Mouse Brain Library and GeneNetwork (presented by Robert Williams). Other resources, such as NeuroData (presented by William R. Gray Roncal), and OpenFMRI, NeuroVault, and Neurosynth (presented by Russell Poldrack) have not been broadly utilized by the neuroscience education community but offer obvious potential.

Finally, William Grisham discussed the iNeuro Project, an NSF-sponsored effort to develop the necessary curriculum for preparing students to handle Big Data. Linda Lanyon further elaborated on the current state and challenges in educating students to deal with Big Data and described some training resources provided by the International Neuroinformatics Coordinating Facility. Neuroinformatics is a subfield of neuroscience that deals with data utilizing analytical tools and computational models. The feasibility of offering neuroinformatics programs at primarily undergraduate institutions was also discussed.

Key words: Big Data; neuroinformatics; educational resources; Allen Brain Atlas; Mouse Brain Library; GeneNetwork   NeuroData;   OpenfMRI;   NeuroVault; Neurosynth; INCF; iNeuro

Beginning in 2005, the annual meeting of the Society for Neuroscience has featured a professional development workshop on different aspects of teaching neuroscience. The 2016 workshop's topic was "Teaching Neuroscience with Big Data." We present here a summary of that workshop, along with an appendix of links to online resources related to the workshop (see Appendix URL List item #1).

Big Data is informally defined as any dataset that will not fit on a laptop, but more generally it includes large repositories and tools that are widely shared. Big Data resources have increasingly become part of the research landscape, and they provide opportunities for training and research in educational settings. Big Data resources extend from the molecular/genomic through whole brain morphology, connectivity, and function. They provide data from large-scale experiments that demand facilities beyond the reach of most educational institutions. Because these resources are available online, educators do not need to invest in large-scale *in situ* hybridization experiments, or produce countless sectioned and stained brains from various strains of recombinant inbred mice, or perform CLARITY/calcium imaging experiments, or have a scanner available to perform structural/functional MRIs. These resources allow instructors and students to pursue questions of genuine interest and in some cases even make novel contributions to neuroscience. Although Big Data resources hold great promise for both research and

education, this promise will not be realized unless students become familiar with the resources and are trained to use them.

Specific resources examined in the 2016 workshop included the Allen Brain Atlas, the Mouse Brain Library and GeneNetwork, the Open Connectome Project, OpenfMRI, Neurosynth and NeuroVault; and Training Space, which is a compendium of neuroinformatics resources provided by the International Neuroinformatics Coordinating Facility (INCF). Finally, there were brief presentations on constructing an educational curriculum in Neuroinformatics.

TEACHING NEUROANATOMY WITH THE ALLEN BRAIN ATLAS

Presented by Terri Gilbert and Joshua C. Brumberg
The Allen Institute for Brain Science is perhaps best known for mapping gene expression in the mouse brain (Lein et al., 2007) by providing high resolution *in situ* hybridization (ISH) images registered into an anatomically annotated three-dimensional space (URL List #2). They have also mapped gene expression across development of brains in three species: mouse, macaque, and human (URL #3-6). In 2010, the Allen Institute began exploring how brain regions are connected, creating a database of cell types and neural coding behind sensory processing. They have also provided several other extensive datasets such as mouse spinal cord ISH, an atlas exploring gene expression in glioblastoma,

and a study of the aging human brain (URL List #7-9).

Obviously, such a rich and varied set of resources can be used in teaching in multiple ways. For example, the Allen Brain Atlas resources can be used to teach neuroanatomy of mice, humans, and macaques, either in isolation or in comparison. The Allen Brain Atlas resources also provide excellent mouse and human reference atlases, which require very little training to use and which have phenomenal in-class utility, allowing students to explore various questions. Some of these applications have been described in previous issues of *JUNE* (e.g., Jenks, 2009; Ramos et al., 2007).
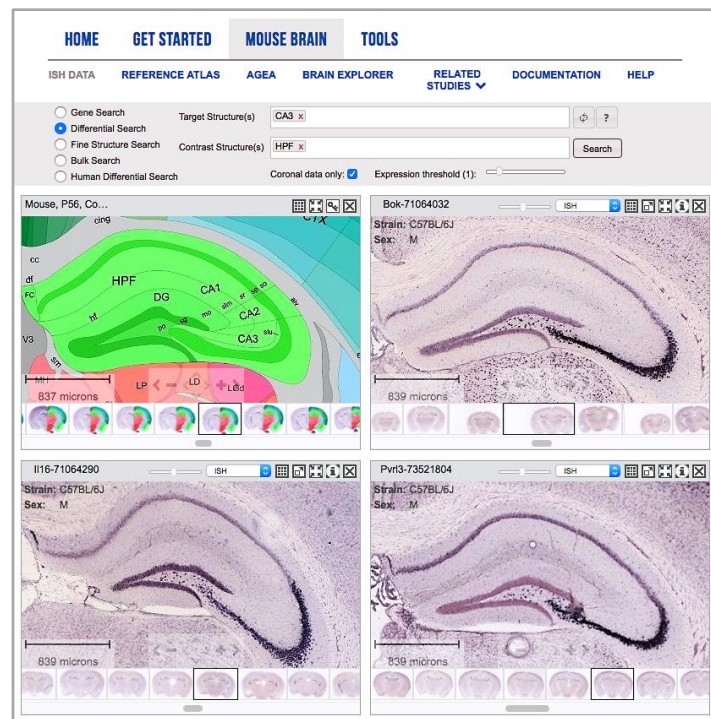

Figure 1. Differential Search Tool. The informatics data processing included in all the gene expression data sets allow for searching for enhanced gene expression in a structure of the brain using Differential Search. For example, in the Allen Mouse Brain Atlas, when choosing CA3 as a target structure and the hippocampal formation (HPF) as a contrast structure, genes with enhanced expression specifically in CA3 relative to the rest of the hippocampal formation are returned.

Areal Segregation. Using the mouse *in situ* hybridization images, students can discover that different brain regions can be segregated by area and/or laminae with regard to the gene expression pattern that they display. The Allen Brain Atlas even provides a background subtraction algorithm to enhance the contrast for segregating anatomical regions. Further, the Allen Brain Atlas gene expression data drive discovery by bridging structural and functional neuroanatomy, allowing the user to generate lists of genes with correlated expression patterns (Figure 1). Indeed, the Allen Brain Atlas has created a Pearson's correlation map based on coronal gene expression that allows users to find voxel-based regions with similar genes expressed within them (Figure 2; URL List #10).

Connectivity. Transgenic tools based on the gene expression atlases have allowed for maps of long range neuronal connections to be collected and visualized across the mouse brain, further adding structural information to the gene expression data (URL List #11). Unique 3D visualization tools also allow the user to see connectivity, gene expression, and structure overlaid in an interactive environment (URL List #12).
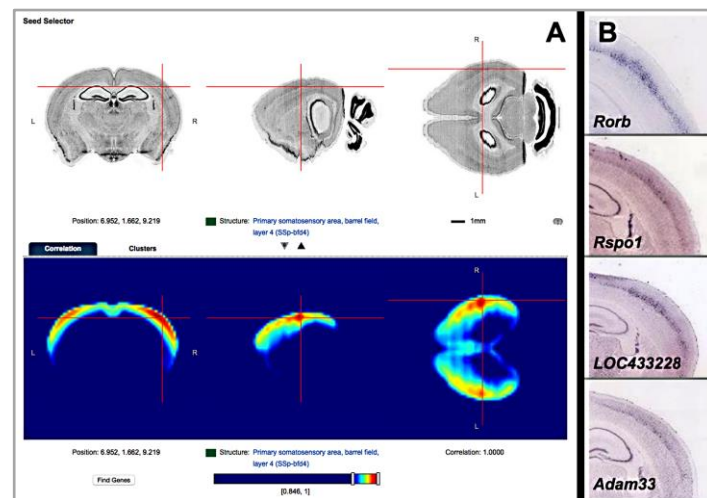

Figure 2. Anatomic Gene Expression Atlas (AGEA). Differential Search can also be applied at the voxel level instead of at the structural level. *(A)* Searching for enhanced gene expression by selecting a voxel via crosshairs in a 3-dimensional space, e.g., centering the crosshairs in layer 4 of the barrel field of the somatosensory area. *(B)* The Find Genes option returns genes with enhanced expression in just that layer (Rorb, Rspo1, LOC433228, Adam33 on the right).

Comparative Anatomy and Development. The Allen Developing Mouse Brain Atlas provides additional resources to enhance teaching neural development, including quantitative investigations such as measuring ventricular volume against measures of cortical plate development. Users can also employ the unique informatic tools in the mouse datasets to synchronize gene expression images both in space and in time during development. Data collection in the Allen Brain Atlas resources involved assaying similar genes across development in mouse, macaque and human. While the data modalities do differ depending upon the species (ISH, microarray, and RNASeq) the data can be compared at similar developmental stages across species.

Cellular Anatomy and Physiology. Again, taking advantage of transgenic tools based on the gene expression mapping (Figure 3), users can begin to correlate gene expression in cell types with cell firing patterns in V1 to various stimuli (see the Allen Cell Types Database, URL List #13; see also the Allen Brain Observatory, URL List #14). For more clinically focused courses, tools for glioblastoma or aging, dementia, and traumatic brain injury that display molecular and transcriptomic characterizations of human cases have been released displaying immunohistochemistry, *in situ* hybridization and RNA-Seq data. Tutorials for navigating and mining the Allen Institute

for Brain Science resources can be found online (URL List #15).

Computational Neuroscience and Integration with Other Big Data Tools. Finally, programmatic access to the data (from image data to microarray/RNA-Seq to fluorescence traces) in the Allen Brain Atlas resources is available for download from the Allen Brain Atlas Application Programing Interface (URL List #16) and the Allen Software Development Kit (AllenSKD; URL List #17). So, users can pursue their own data mining on the Allen Institute resources.

The resources provided by the Allen Institute also can be used in conjunction with other big data tools (see below) such as GeneNetwork and the UC Santa Cruz Genome browser, which adds value to a bioinformatics unit by tying together Quantitative Trait Loci data or structural data with regional gene expression (see Grisham et al., 2012).
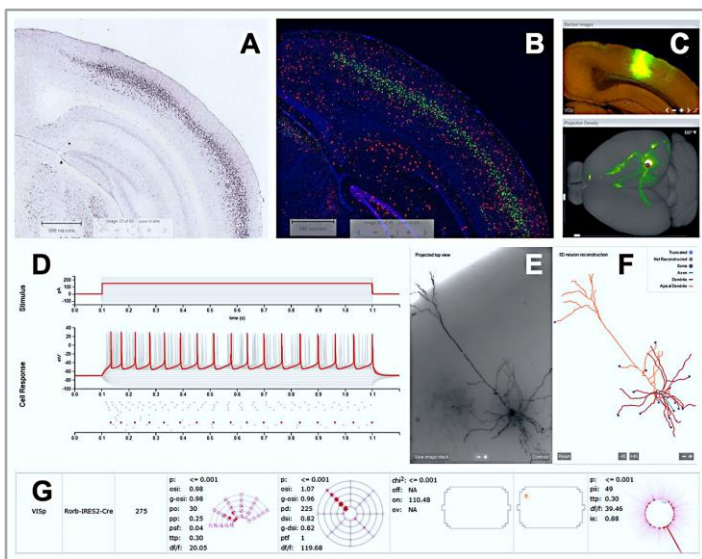

Figure 3. Genetic tools to probe cell class and function. *In situ* hybridization (ISH) gene expression patterns of Rorb from the Allen Mouse Brain Atlas *(A)* led to the creation of the transgenic mouse line Rorb-IRES-Cre (green label in *B*). This transgenic line was used to probe connectivity (*C*) of the cell classes (Allen Mouse Brain Connectivity Atlas), electrophysiological (*D*) and morphological characterization (*E, F*) at the single cell level (Allen Cell Types Database), as well as understanding the calcium signaling of these cells to visual stimuli (*G*) as assayed in the Allen Brain Observatory.

MOUSE BRAIN LIBRARY (MBL) AND GENENETWORK

Presented by Robert W. Williams

The Mouse Brain Library consists of Nissl-stained images of sections of brains of mice from many genetically characterized strains, including various inbred and recombinant inbred strains with individuals of both sexes (Figure 4; URL List #18). The aim of this library is to understand normal genetic variants (rather than mutations) that generate differences in cell populations and cell phenotypes with a wide range of behavioral and functional traits that have also been measured with these genotypes.

The collection presently consists of images from approximately 800 brains and numerical data from over 8000 mice. The MBL can be searched by metadata such as strain, age, sex, body or brain weight, which were purposely varied in the samples. Such searches allow students to explore differences among brains, for example, sex differences (Grisham et al., 2016).
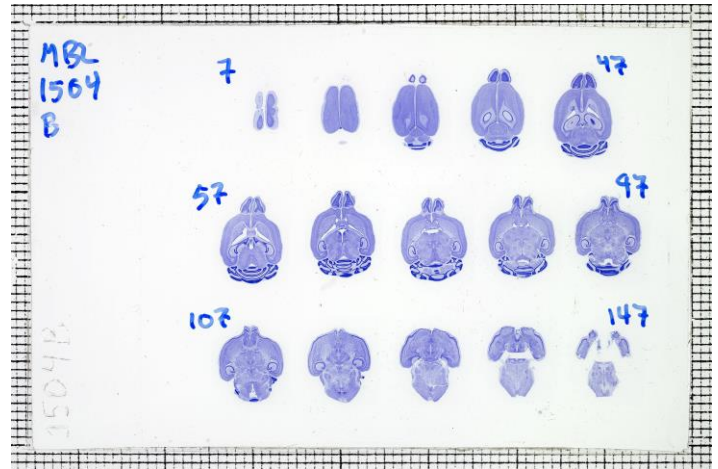

Figure 4. Mouse Brain Library (MBL) sections. High-resolution image of a BXD mouse brain (1504B) cut in the horizontal plane at 30μm, with every 10th section mounted and stained with cresyl violet.

The MBL can be used in conjunction with GeneNetwork (URL List #19) and other Big Data tools to create a bioinformatics unit (see Grisham et al., 2012, for an in-depth description of using these tools in teaching; URL List #41). Students can utilize the MBL images from various inbred strains and quantify a given brain phenotype. The recombinant inbred strains provide differences in genotype that can then be correlated with differences in phenotypic measures via tools in GeneNetwork. These tools can perform a Quantitative Trait Locus (QTL) analysis (Figure 5), which attempts to track down a chromosome region likely to be contributing to heritable variation in the phenotype. Thus, a causal gateway between loci in the genome and variance in phenotype is provided (see Mulligan et al., 2017, for a detailed primer in using GeneNetwork within a computer lab setting). Students can zoom in on the chromosome(s) and using the SNP density to guide them, they can link out to potential candidate genes displayed in the UCSC Genome Browser, thereby finding one that shows high gene expression in the region of interest (URL List #20). Students can then link to the Allen Brain Atlas and explore the *in situ* gene expression data of individual cells in the region. Finally, students can link to big data repositories such as Gene and PubMed for more information about given genes (URL List #21-22).

The resources of the MBL and GeneNetwork have been successfully used in a jamboree-model short course on neuroinformatics for the International Neuroinformatics Coordinating Facility in which groups worked intensively for a week learning and utilizing neuroinformatics tools and resources (Overall et al., 2015; Heimel et al., 2015).
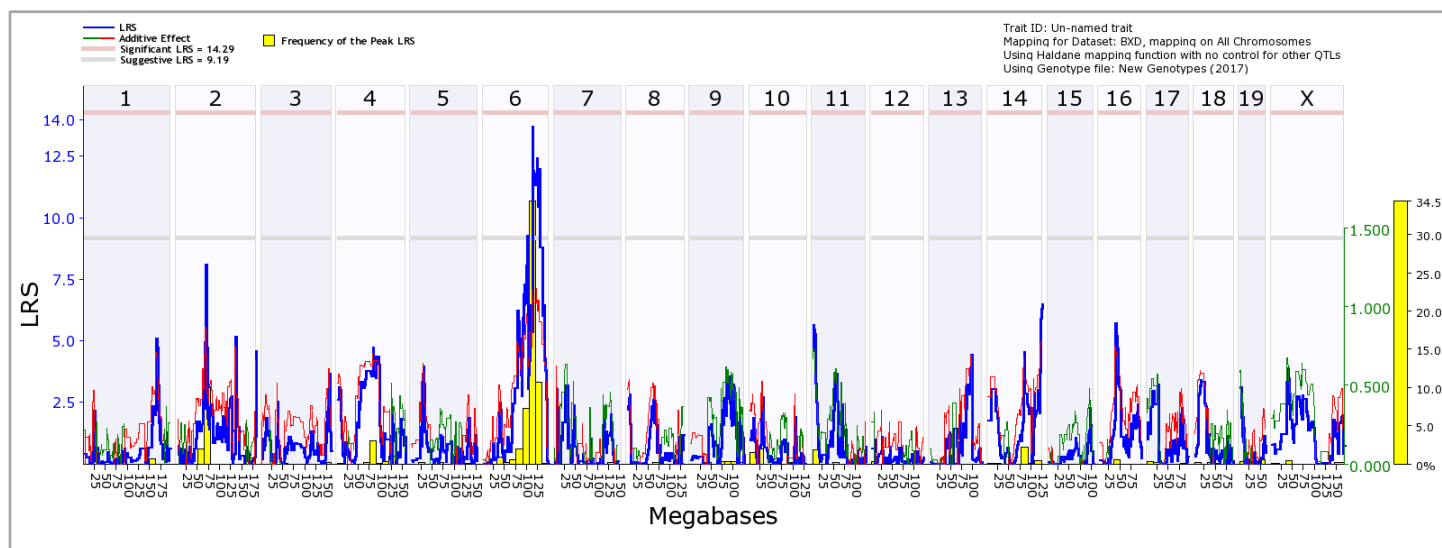
Figure 5. GeneNetwork Quantitative Trait Locus (QTL) mapping model. Average residual variance, based on olfactory bulb size, for various BXD recombinant inbred strains of mice are entered into GeneNetwork and a whole genome map is generated. Here we see that a region of chromosome 6 (chromosome number indicated at the top of the graph) is implicated in olfactory bulb size because the Likelihood Ratio Statistic (LRS) exceeds the suggested criterion (gray line).

Participants were divided into small focus groups of four to five and addressed such topics as "Neural Development," "Addiction and Impulsivity," "Brain Disorders," "Synapse and Plasticity," "Neurodegeneration," and "Adult Neurogenesis." These teams then used tools such as the Allen Mouse Brain Atlas, the BrainSpan Atlas of the Developing Human Brain, and GeneNetwork, among others, to delve into their respective topics. The participants produced six draft manuscripts during the intensive one-week workshop, followed by several months of editing and revision, which resulted in six manuscripts that passed peer review. A full semester jamboree model could serve as a framework for undergraduate education. Faculty could construct courses using the available neuroinformatics tools to engage students in authentic research questions as per the recommendations of Vision and Change (Ledbetter, 2012).

NEURODATA

Presented by William R. Gray Roncal (*Johns Hopkins University*)

NeuroData is a newer and exciting resource that leverages cloud-computing technologies to enable large-scale neurodata storing, exploring, analyzing, and modeling (URL List #23). Its authors have created a "computational ecosystem" for enabling petascale neuroscience. In particular, the NeuroData team is examining neural circuits at a variety of levels and from multiple perspectives. The interests of NeuroData span the gamut of spatiotemporal scales of analysis, ranging from nanoscale (serial electron microscopy) to microscale (array tomography, CLARITY, and calcium imaging) to macroscale (SPECT and multimodal magnetic resonance imaging). In many ways, exploring their website is like taking a tour of what neuroinformatics will look like in the future. Both the data and code of NeuroData are open source and public.

Like GeneNetwork, NeuroData also provides tools for data analyses such as CLARITY visualization (URL List #24). Tools such as NeuroData's MRI Graphs (NDMG) pipeline can combine diffusion MRI with structural MRI images to determine the connectome in a given subject (URL List #25). Some of the tools, such as FlashX can analyze data at mind-boggling scales such as billion-node graphs (URL List #26). Other tools, like Multiscale Generalized Correlation can examine variables ranging from the molecular level to those that might determine behavior, as well as levels in between, and look at the relationships among these variables (URL List #27).

The tools of NeuroData come with documentation and often GitHub repositories, and are open source and fairly accessible if approached with an "adaptable mindset," including a sense of curiosity mixed with tenacity. Two tools
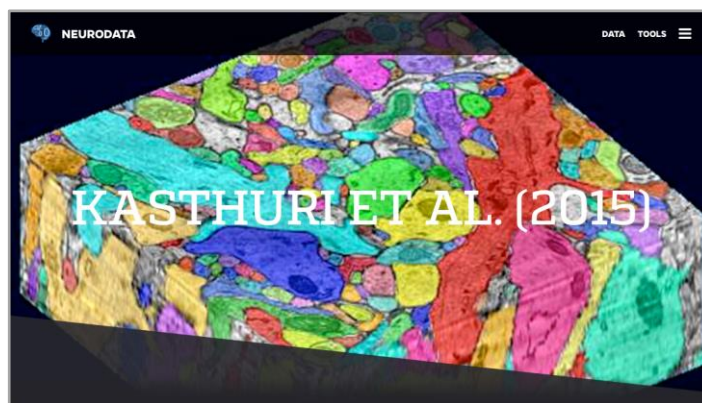

Figure 6. Multi-scale cortical reconstruction. Multi-scale reconstruction of a mouse cortex by Kasthuri et al. (2015) using detailed electron microscopy. These data are available via NeuroData's NDIO module.

of NeuroData are relatively easy to use: NeuroData Input-Output (NDIO) and NeuroData PARSE. NDIO can be installed in two minutes and allows pulling data from sources such as Kasthuri et al., 2015 (Figure 6), which has extremely detailed electron microscopy (EM) data on the mouse cortex. NeuroData PARSE is a repository holding code needed to train, evaluate, and deploy code for parsing volumes of NeuroData images.

Although the domain of Big Data is complex, connectomics may be an excellent topic for teaching Big Data, and appropriate tools can accelerate the learning process. Connectomics has been successfully taught as an undergraduate course/extended workshop, including to students who initially did not know command line computing.

The resources provided by NeuroData are quite impressive, but utilization of the tools may require instructors to take a deeper dive to understand them well enough to employ them in undergraduate classes. Nonetheless, the dive may well be worth it.

OPENFMRI, NEUROVAULT, AND NEUROSYNTH

Presented by Russell Poldrack *(Department of Psychology, Stanford University)*
This set of tools and resources can be seen as a model for an eco-system of open neuroscience data that would not only encourage investigators to contribute data but also enhance discoverability and sustainability of resources (Wiener et al., 2016). This eco system would not only include standards for shared data but also would encourage tool sharing, including virtual machines (Eglen et al., 2017). Accordingly, this set of resources and tools have a GitHub repository for every aspect of the workflows, including analyses and notebooks.

Several exciting neuroinformatic tools were introduced, including OpenfMRI, NeuroVault, and Neurosynth (URL List #28-30). The range of resources will prove to be of substantial value to educators and will allow for many levels of involvement from basic pointing-and-clicking to writing scripts for analyses. These resources extend from a ready-to-use tool that can enhance students' understanding of brain function to other tools that might require additional skills but will nonetheless be very useful.

OpenfMRI is a web resource that provides raw fMRI data from various experiments, making these data sets available for analyses and opening them for questions not posed in their initial publications. Most of these datasets are from published experiments, which would allow novices to check their analyses against an expert's. OpenfMRI employs the Brain Imaging Data Structure (BIDS) format, which provides a standard for organizing the shared data and making them immediately usable by other researchers. Using OpenfMRI requires knowledge of other MRI analysis software such as the FMRIB Software Library (FSL; URL List #31). FSL is available for free, however, so the potential utility of resources such as OpenfMRI is considerable (Figure 7). FSL is a command-line package and requires some learning, but hopefully such skills will become more widespread and will propagate across the undergraduate
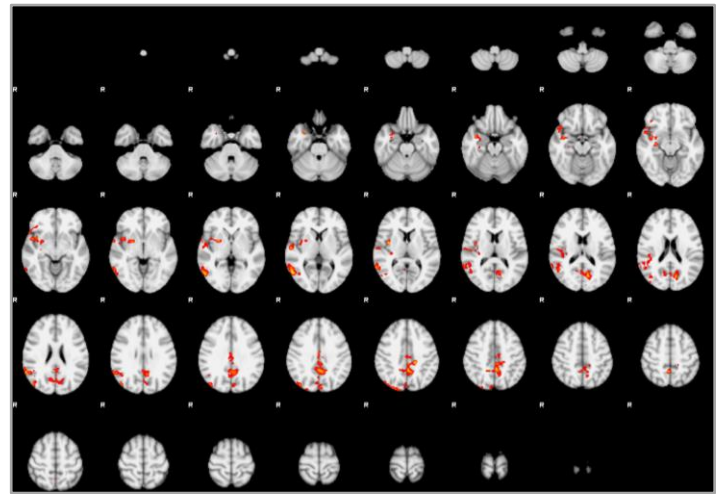

Figure 7. OpenfMRI dataset analysis. Images generated by a UCLA undergraduate using an OpenfMRI dataset and FSL.

Figure 8. NeuroVault Collections publication. An article by Harper et al., 2017, including interactive results, are hosted publicly on NeuroVault. This particular figure illustrates voxels that are likely to be altered in patients with early-onset Alzheimer's Disease compared to controls.

teaching/learning community.

Almost all the datasets in OpenfMRI have a corresponding entry in NeuroVault, an online repository for brain statistical maps. Notably, these are unthresholded statistical maps, rather than thresholded activation maps, so they provide more detail about the statistical findings of the study such as differences between groups (Figure 8; Harper et al., 2017) or independent component analyses mapped onto a standard brain (e.g., Bethlehem et al., 2017; see also McKeown et al., 2003, for a discussion of independent component analysis (ICA) in fMRI). NeuroVault can not only start discussions and serve as the starting points of projects on the anatomical underpinnings of neural processes, but also it can serve as a starting point for educating students about sophisticated statistical analyses. Mumford Brain Stats, a great resource spanning the range from basic stats to sophisticated analyses, and NeuroStars, a resource where experienced users answer online questions mostly

dealing with various MRI analyses, were also mentioned (URL List #32-33).

Neurosynth is a site that automatically extracts coordinates from published papers via text mining and displays the results mapped onto a standard brain (Figure 9), showing regions that are activated in given studies. For example, Figure 9 displays the result of searching the term "spatial" using both forward inference (studies that manipulated a psychological process and identified a corresponding localized effect on brain activity) and reverse inference (studies that started with patterns of activation to infer the mental processes involved; see Poldrack, 2011, for a discussion of forward and reverse inference). Regardless of the path of inference, Neurosynth allows students to explore structure-function relationships within the brain. Further, as with NeuroVault, students are able to pull up meaningful brain maps instantly without requiring any computer skills beyond point-and-click. The gene expression data from the Allen Brain Atlas has been incorporated into Neurosynth so that regions of gene expression can be visualized on a human brain. Neurosynth also allows students to retrieve the list of papers from which the data was mined.

Figure 9. Neurosynth Meta-Analyses. An automated meta-analysis of 1157 studies in the Neurosynth library. Blue represents forward inference, and red represents reverse inference for the searchword "spatial."

TEACHING STUDENTS ABOUT NEUROINFORMATICS—INTERNATIONAL NEUROINFORMATICS COORDINATING FACILITY (INCF) AND THE INEURO PROJECT

Presented by Linda Lanyon (*INCF*) and William E. Grisham (*iNeuro*)

Neuroscience is an increasingly data-intense field. Big Data training from the undergraduate through postgraduate levels is going to be important in preparing neuroscientists for the future of the field. Clearly, the next generation of neuroscientists must be equipped with neuroinformatics skills. Further, even if neuroscientists are not exclusively focused on neuroinformatics or computational neuroscience, it would behoove all neuroscientists to at least be conversant with aspects of this field. Current neuroscience faculty may not be well-equipped to deliver neuroinformatics education because this did not form part of their own training path. Hence, faculty development and instructor education/training is crucial.

The iNeuro Project was an NSF-funded effort that explored where we are now and where we need to go in neuroinformatics training (see the iNeuro Project Workshop Report; URL List #34-35). The iNeuro Project addressed why we need neuroinformatics programs, what curriculum these programs should contain, and how we should go about constructing these programs. Curricula should be driven by the desired skills sets that educators wish to instill in their students. In the case of neuroinformatics training, iNeuro Workshop participants decided that trainees should have a neuroscience background that included experimental design with wet lab skills, technical/computing/analytic skills, information science/informatics/data science skills along with quantitative data analysis, soft skills such as management and team building, as well as a knowledge of
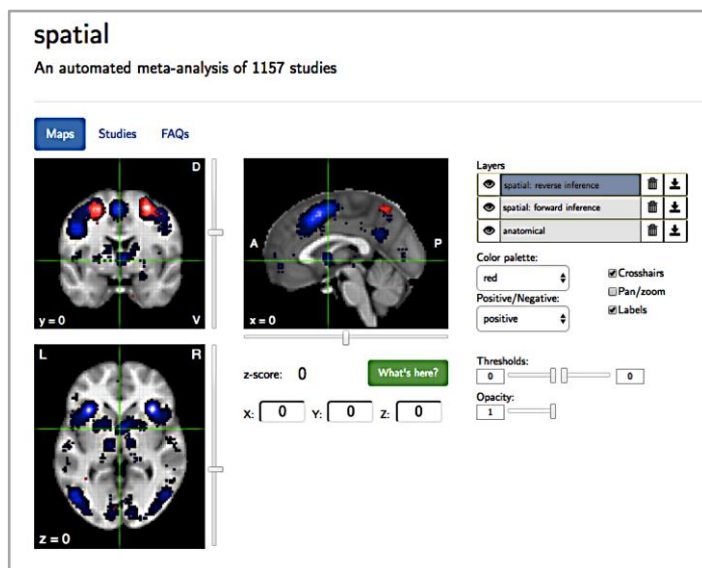
ethical considerations. Possible curricula are outlined in the iNeuro Workshop Report (URL List #35).

There is an urgent, global demand for integration of neuroinformatics in mainstream neuroscience education. There are few programs worldwide that focus on neuroinformatics training, and presently none exist in the United States (Grisham et al., 2016; iNeuro Project Workshop Report, URL List #35). Current programs that focus exclusively on neuroinformatics training exist in Europe; there are graduate programs in Edinburgh and Newcastle, UK, and KTH Royal Institute of Technology, Sweden. There is presently one undergraduate program at the University of Warsaw, Poland. In addition, the INCF runs an annual two-day course with introductory topics (URL List #36-37).

Extant online resources for neuroinformatics training are also sparse and disparate. Issues include difficulties in discoverability, lack of curricula and training pathways for beginners, gaps in content, and the lack of training support and quality rating of materials. INCF currently supports a YouTube channel, which contains a number of training videos along with supporting materials (URL List #38). INCF is working with partner organizations to identify and synthesize existing global resources and, in the near future, INCF will roll out Training Space, an online hub of resources for education about informatics in neuroscience (Figure 10). Training Space will integrate existing resources and encourage further community-driven content creation, providing global access to multimedia curricula and content for neuroinformatics training and education. Among those resources will be the NIH Big Data to Knowledge (BD2K) Training Coordinating Center (TCC) education initiative, which features the Educational Resource Discovery Index (ERuDIte) database (URL List #39). This index now contains information on scores of resources related to
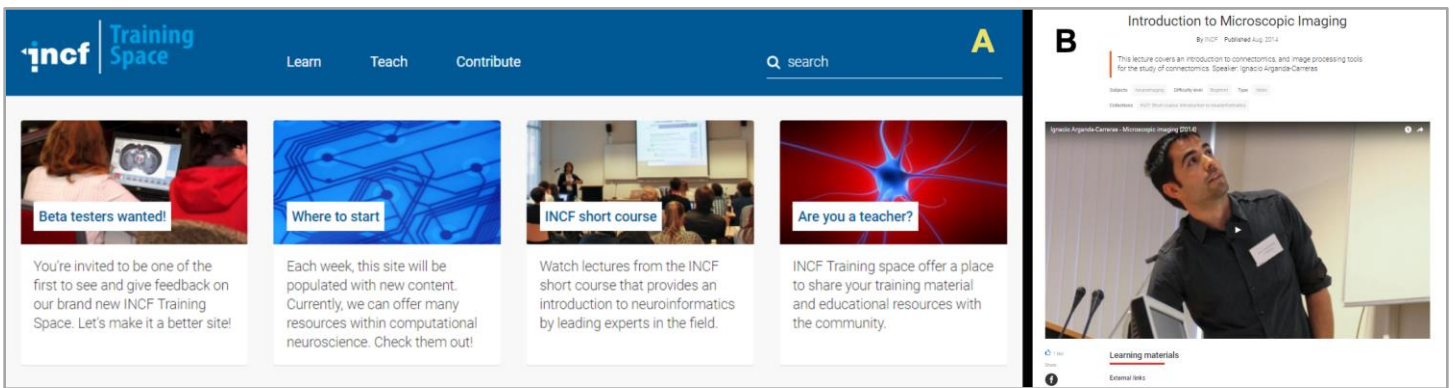
Figure 10. Training Space. Developed by INCF and partners, Training Space is an online hub that links neuroinformatics training resources. *(A)* Landing page. *(B)* Sample lecture content.
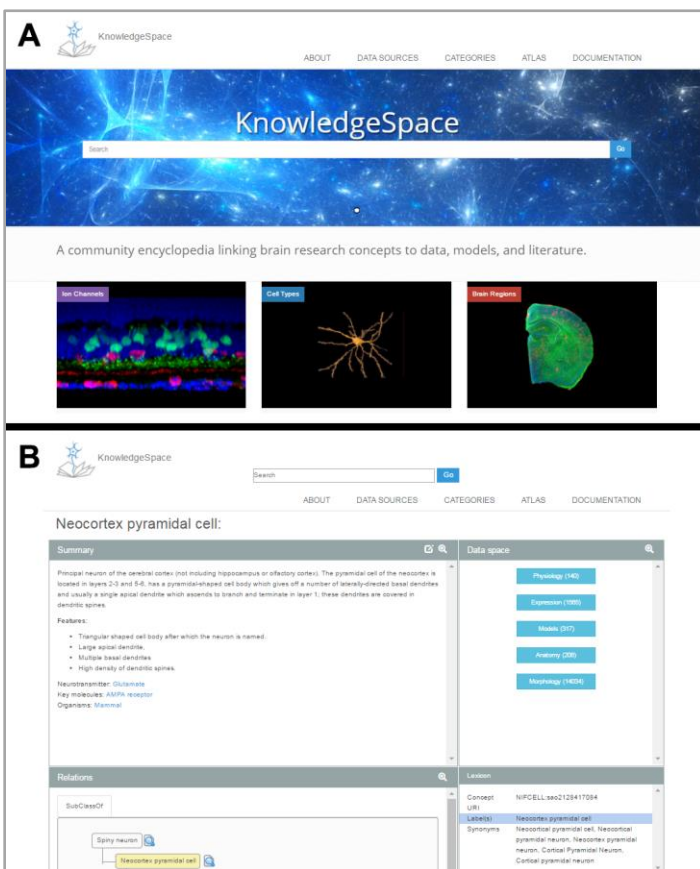

Figure 11. KnowledgeSpace. Developed by INCF and partners, KnowledgeSpace links brain research concepts with data, models, and literature. *(A)* Landing page. *(B)* Information on a Neocortex pyramidal cell.

neuroinformatics data science training. The KnowledgeSpace encyclopedia for neuroscience (Figure 11; URL List #40) is a further resource being linked to Training Space. KnowledgeSpace is currently in pilot form developed in partnership by INCF, the Neuroscience Information Framework, and the Human Brain Project. It aims to be a community-driven encyclopedia linking brain research concepts to data, computational models, and research literature.

Finally, one question that would be of interest to *JUNE* readers was raised in the iNeuro Workshop: can Big Data neuroinformatics curricula be incorporated into a small liberal arts college or are such curricula the purview of large-scale research universities? Recently the University of Washington (Seattle, WA) and Northwestern University (Evanston, IL) have announced new Big Data-focused educational programs. Some participants at the iNeuro Workshop argued that Big Data neuroinformatics programs would be better situated at large-scale research universities since they have a greater depth and breadth of faculty resources.

Nonetheless, the University of Puerto Rico and primarily undergraduate institutions (PUIs) such as the California State Universities at Northridge and Monterey Bay are developing Big Data programs that have a unique focus on the inclusion of underrepresented minorities and first-generation students (Van Horn, 2016). So, programs focusing on Big Data in neuroscience/neuroinformatics at PUIs do not seem far-fetched. Further, Big Data resources are available and their utility in undergraduate education is undeniable. In future years, we expect neuroscience faculty to increasingly embrace these resources in their teaching, providing their students with rich arrays of data resources and opportunities for genuine inquiry in line with the principles advanced by *Vision and Change* (Anderson et al., 2013).

GETTING STARTED

Utilizing some Big Data resources requires more instructor preparation than others. Some require either knowing or acquiring skills such as coding in Python. Other tools, however, are easy to adopt/adapt for teaching purposes. Currently, tools that are relatively easy to use are NeuroVault and Neurosynth, which allow students to explore results of studies on humans with no specialized training. Employing NeuroVault and Neurosynth with the interactive Allen Human Brain Atlas as a reference will undoubtedly enhance students' understanding of the neuroinformatics outputs.

A teaching module in neuroinformatics that is fairly well developed and designed to be plug-and-play has been outlined above (in the section titled **MOUSE BRAIN**

LIBRARY (MBL) AND GENENETWORK; see also Grisham et al., 2012). This neuroinformatics/bioinformatics module (URL List #41) integrates several Big Data resources and tools including the Mouse Brain Library, GeneNetwork, UCSC Genome Browser, Gene, PubMed, and the Allen Mouse Brain Atlas. A separate module utilizing the Allen Brain Atlas as a stand-alone teaching tool to reveal the cytoarchitecture, cellular diversity, and gene expression profiles of the brain has also been well described (Ramos et al., 2007). Finally, a biocuration undergraduate research program for constructing databases of Amyotrophic Lateral Sclerosis (ALS) transgenic mouse models and clinical data has recently been described in a *JUNE* article (Mitchell et al., 2015).

Low-cost resources that might better prepare faculty to utilize Big Data can be found in the INCF YouTube channel training videos, which will proliferate following the launch of INCF's Training Space. Additionally, there are some online courses on MRI imaging via Coursera and edX, and other sources, many of which are free (URL List #42-44).

APPENDIX: LIST OF URLS

INTRODUCTION
(1) 2016 Society for Neuroscience Teaching Workshop
http://funfaculty.org/teaching_workshops/2016/index.html

TEACHING NEUROANATOMY WITH THE ALLEN BRAIN ATLAS
(2) Allen Mouse Brain Atlas
http://mouse.brain-map.org/
(3) Allen Developing Mouse Brain Atlas
http://developingmouse.brain-map.org/
(4) NIH Blueprint Non-Human Primate (NHP) Atlas
http://blueprintnhpatlas.org/
(5) Allen Human Brain Atlas
http://human.brain-map.org/
(6) BrainSpan Atlas of the Developing Human Brain
http://brainspan.org/
(7) Allen Mouse Spinal Cord Atlas
http://mousespinal.brain-map.org/
(8) IVY Glioblastoma Atlas Project (GAP)
http://glioblastoma.alleninstitute.org/
(9) Aging, Dementia, and TBI Study
http://aging.brain-map.org/
(10) Allen Mouse Brain Atlas Anatomic Gene Expression Atlas (AGEA)
http://mouse.brain-map.org/agea
(11) Allen Mouse Brain Connectivity Atlas
http://connectivity.brain-map.org/
(12) Brain Explorer 2
http://connectivity.brain-map.org/static/brainexplorer
(13) Allen Cell Types Database
http://celltypes.brain-map.org/
(14) Allen Brain Observatory
http://observatory.brain-map.org/visualcoding/
(15) Allen Brain Atlas Tutorials
http://www.brain-map.org/tutorials/index.html
(16) Allen Brain Atlas Application Programming Interface (API)
http://brain-map.org/api/index.html
(17) Allen Software Development Kit (AllenSKD)
http://alleninstitute.github.io/AllenSDK/

MOUSE BRAIN LIBRARY (MBL) AND GENENETWORK
(18) The Mouse Brain Library
http://www.mbl.org/
(19) GeneNetwork
http://www.genenetwork.org/
(20) UCSC Genome Browser
https://genome.ucsc.edu/
(21) NCBI Gene
https://www.ncbi.nlm.nih.gov/gene
(22) NCBI PubMed
https://www.ncbi.nlm.nih.gov/pubmed/

NEURODATA
(23) NeuroData
https://neurodata.io/
(24) Chung Lab CLARITY
http://www.chunglabresources.com/clarity/
(25) GitHub: NeuroData's MRI Graphs (NDMG)
https://github.com/neurodata/ndmg
(26) NeuroData FlashX
https://neurodata.io/tools/FlashX/
(27) NeuroData Multiscale Generalized Correlation
https://neurodata.io/tools/MGC/

OPENFMRI, NEUROVAULT, AND NEUROSYNTH
(28) OpenfMRI
https://openfmri.org/
(29) NeuroVault
http://neurovault.org/
(30) Neurosynth
http://neurosynth.org/
(31) FMRIB Software Library (FSL)
https://fsl.fmrib.ox.ac.uk/fsl/fslwiki
(32) Mumford Brain Stats
http://mumfordbrainstats.tumblr.com/
(33) NeuroStars
https://neurostars.org/

TEACHING STUDENTS ABOUT NEUROINFORMATICS—INTERNATIONAL NEUROINFORMATICS COORDINATING FACILITY (INCF) AND THE INEURO PROJECT
(34) iNeuro Project
https://mdcune.psych.ucla.edu/ineuro
(35) iNeuro Project Workshop Report
https://mdcune.psych.ucla.edu/ineuro/reports/iNeuro_WorkshopReport_v20170123.pdf
(36) International Neuroinformatics Coordinating Facility (INCF)
https://www.incf.org/
(37) INCF Congress of Neuroinformatics
https://www.incf.org/collaborate/incf-congress-of-neuroinformatics
(38) YouTube: INCF
https://www.youtube.com/user/INCForg
(39) NIH Big Data to Knowledge (BD2K) Training Coordinating Center (TCC)
https://bigdatau.ini.usc.edu/
(40) KnowledgeSpace
https://knowledge-space.org/

GETTING STARTED
(41) MDCUNE Bioinformatics/Neuroinformatics
https://mdcune.psych.ucla.edu/modules/bioinformatics

(42) Coursera
https://www.coursera.org/
(43) edX
https://www.edx.org/
(44) Neuroscience News: Free Neuroscience MOOCs
https://neurosciencenews.com/free-neuroscience-moocs/

REFERENCES

Anderson CW, Bauerle C, DePass A, Lynn D, O'Connor C, Singer S, Withers M, Anderson CW, Donovan S, Drew S, Ebert-May D, Gross L, Hoskins SG, Labov J, Lopatto D, McClatchey W, Varma-Nelson P, Pelaez N, Poston M, Tanner K, Wessner D, White H, Wood W, Wubah D (2013) Vision and change in undergraduate biology education: a call to action. Brewer CA, Smith D (Eds.). Retrieved from http://visionandchange.org/files/2013/11/aaas-VISchange-web1113.pdf.

Bethlehem RAI, Lombardo MV, Lai MC, Auyeung B, Crockford SK, Deakin J, Soubramanian S, Sule A, Kundu P, Voon V, Baron-Cohen S (2017) Intranasal oxytocin enhances intrinsic corticostriatal functional connectivity in women. Transl Psychiatry 2017 7(4):e1099. doi: 10.1038/tp.2017.72.

Eglen SJ, Marwick B, Halchenko YO, Hanke M, Sufi S, Gleeson P, Silver RA, Davison AP, Lanyon L, Abrams M, Wachtler T, Willshaw DJ, Pouzat C, Poline JB (2017) Toward standard practices for sharing computer code and programs in neuroscience. Nat Neurosci 20:770-773. doi: 10.1038/nn.4550.

Grisham W, Korey CA, Schottler NA, McCauley LB, Beatty J (2012) Teaching neuroinformatics with an emphasis on quantitative locus analysis. J Undergrad Neurosci Educ 11(1):A119-A125.

Grisham W, Lom B, Lanyon L, Ramos RL (2016) Proposed training to meet challenges of large-scale data in neuroscience. Front Neuroinform 10:28. doi: 10.3389/fninf.2016.00028.

Grisham WE, Tsai C, Inouye K (2016, November) Sex differences in the dentate gyrus and overall shape of the hippocampus in DBA mice. Neuroscience 2016 Abstracts 179.01/III37. Abstract and poster presented at the Society for Neuroscience Annual Meeting, San Diego, CA.

Harper L, Bouwman F, Burton EJ, Barkhof F, Scheltens P, O'Brien JT, Fox NC, Ridgway GR, Schott JM (2017) Patterns of atrophy in pathologically confirmed dementias: a voxelwise analysis. J Neurol Neurosurg Psychiatry 0:1-9. doi: 10.1136/jnnp-2016-314978.

Heimel JA, Overall RW, Williams RW (2015) Workshop report: INCF short course on neuroinformatics, neurogenomics, and brain disease, 14–21 September 2013. Front Neurosci 9:31. doi: 10.3389/fnins.2015.00031.

iNeuro Project Workshop Report: Preparing a workforce to meet the challenges of large-scale neuroscience data; Producing curricula and resources for large-scale neuroscience data analysis. (2017). Retrieved from https://mdcune.psych.ucla.edu/ineuro/reports/iNeuro_WorkshopReport_v20170123.pdf.

Jenks BG (2009) A self-study tutorial using the Allen Brain Explorer and Brain Atlas to teach concepts of mammalian neuroanatomy and brain function. J Undergrad Neurosci Educ 8(1):A21-A25.

Kasthuri N, Hayworth KJ, Berger DR, Schalek RL, Conchello JA, Knowles-Barley S, Lee D, Vázquez-Reina A, Kaynig V, Jones TR, Roberts M, Morgan JL, Tapia JC, Seung HS, Roncal WG, Vogelstein JT, Burns R, Sussman DL, Priebe CE, Pfister H, Lichtman JW (2015) Saturated Reconstruction of a Volume of Neocortex. Cell 162(3):648-61. doi: 10.1016/j.cell.2015.06.054.

Ledbetter MLS (2012) Vision and change in undergraduate biology education: a call to action presentation to Faculty for Undergraduate Neuroscience, July 2011. J Undergrad Neurosci Educ 11(1):A22-A26.

Lein ES, Hawrylycz MJ, Ao N, Ayres M, Bensinger A, Bernard A, Boe AF, Boguski MS, Brockway KS, Byrnes EJ, Chen L, Chen L, Chen TM, Chin MC, Chong J, Crook BE, Czaplinska A, Dang CN, Datta S, Dee NR, Desaki AL, Desta T, Diep E, Dolbeare TA, Donelan MJ, Dong HW, Dougherty JG, Duncan BJ, Ebbert AJ, Eichele G, Estin LK, Faber C, Facer BA, Fields R, Fischer SR, Fliss TP, Frensley C, Gates SN, Glattfelder KJ, Halverson KR, Hart MR, Hohmann JG, Howell MP, Jeung DP, Johnson RA, Karr PT, Kawal R, Kidney JM, Knapik RH, Kuan CL, Lake JH, Laramee AR, Larsen KD, Lau C, Lemon TA, Liang AJ, Liu Y, Luong LT, Michaels J, Morgan JJ, Morgan RJ, Mortrud MT, Mosqueda NF, Ng LL, Ng R, Orta GJ, Overly CC, Pak TH, Parry SE, Pathak SD, Pearson OC, Puchalski RB, Riley ZL, Rockett HR, Rowland SA, Royall JJ, Ruiz MJ, Sarno NR, Schaffnit K, Shapovalova NV, Sivisay T, Slaughterbeck CR, Smith SC, Smith KA, Smith BI, Sodt AJ, Stewart NN, Stumpf KR, Sunkin SM, Sutram M, Tam A, Teemer CD, Thaller C, Thompson CL, Varnam LR, Visel A, Whitlock RM, Wohnoutka PE, Wolkey CK, Wong VY, Wood M, Yaylaoglu MB, Young RC, Youngstrom BL, Yuan XF, Zhang B, Zwingman TA, Jones AR (2007) Genome-wide atlas of gene expression in the adult mouse brain. Nature 445(7124):168-176.

McKeown MJ, Hansen LK, Sejnowsk TJ (2003) Independent component analysis of functional MRI: what is signal and what is noise? Curr Opin Neurobiol 13(5):620-629.

Mitchell CS, Cates A, Kim RB, Hollinger SK (2015) Undergraduate biocuration: developing tomorrow's researchers while mining today's data. J Undergrad Neurosci Educ 14(1):A56-65.

Mulligan MK, Mozhui K, Prins P, Williams RW (2017) GeneNetwork: a toolbox for systems genetics. Methods Mol Biol 1488:75-120.

Overall RW, Williams RW, Heimel JA (2015) Collaborative mining of public data resources in neuroinformatics. Front Neurosci 9:90. doi: 10.3389/fnins.2015.00090.

Poldrack RA (2011) Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. Neuron 72(5):692-697. doi: 10.1016/j.neuron.2011.11.001.

Ramos RL, Smith PT, Brumberg JC (2007) Novel in silico method for teaching cytoarchitecture, cellular diversity, and gene expression in the mammalian brain. J Undergrad Neurosci Educ 6(1):A8-A13.

Van Horn JD (2016) Opinion: big data biomedicine offers big higher education opportunities. Proc Natl Acad Sci U S A 113(23):6322-6324. doi: 10.1073/pnas.1607582113.

Wiener M, Sommer, FT, Ives ZG, Poldrack RA, Litt B (2016) Enabling an Open Data Ecosystem for the Neurosciences. Neuron 92(3):617-621. doi: 10.1016/j.neuron.2016.10.037.

Address correspondence to: William Grisham, Ph.D., Department of Psychology, P.O. Box 951563, University of California, Los Angeles; Los Angeles, CA 90095-1563. Email: dr.billgrisham@gmail.com.