Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Neuroinformatics. Author manuscript; available in PMC 2024 December 10.

Published in final edited form as:

Neuroinformatics. 2024 October ; 22(4): 647–655. doi:10.1007/s12021-024-09666-6.

Hands-on neuroinformatics education at the crossroads of online and in-person: lessons learned from NeuroHackademy

Ariel Rokem[1,2,†], **Noah C. Benson**[2,†]

[1]Department of Psychology, University of Washington, 119 Guthrie Hall, Seattle, 98195, Washington, USA

[2]eScience Institute, University of Washington, 3910 15th Ave NE, Seattle, 98195, Washington, Country

Abstract

NeuroHackademy (https://neurohackademy.org) is a two-week event designed to train early-career neuroscience researchers in data science methods and their application to neuroimaging. The event seeks to bridge the big data skills gap by introducing participants to data science methods and skills that are often ignored in traditional curricula. Such skills are needed for the analysis and interpretation of the kinds of large and complex datasets that have become increasingly important to neuroimaging research due to concerted data collection efforts. In 2020, the event rapidly pivoted from an in-person event to an online event that included hundreds of participants from all over the world. This experience and those of the participants substantially changed our valuation of large online-accessible events. In subsequent events held in 2022 and 2023, we have developed a "hybrid" format that includes both online and in-person participants. We discuss the technical and sociotechnical elements of hybrid events and discuss some of the lessons we have learned while organizing them. We emphasize in particular the role that these events can play in creating a global and inclusive community of practice in the intersection of neuroimaging and data science.

Keywords

Hackathons; Remote learning; Data science; Programming; Community of Practice

1 Introduction

1.1 Human neuroscience in the age of large open datasets

The study of the human brain has entered a new and exciting era of big data. Concerted data collection efforts such as the Human Connectome Project [1, 2], the Adolescent Brain Cognitive Development (ABCD) study [3], the 1000 Functional Connectomes Project (FCP) and its successor, the International Neuroimaging Datasharing Initiative (INDI)[4], and many other similar projects, have recently given the human neuroscience research

Corresponding authors: arokem@uw.edu; nben@uw.edu.
†These authors contributed equally to this work.

Competing Interest Declaration. The authors declare that there are no conflicts of interest regarding this article.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

community access to datasets whose scales and scopes were infeasible even just a decade ago.

Thanks to their scales, and to the efforts made to sample individuals with a variety of characteristics and susceptibilities to a range of neurodevelopmental and mental health disorders, these datasets have substantial scientific potential both for basic and clinical research. However, they also pose a host of new challenges for researchers who are generating, processing, accessing, analyzing, and trying to understand these data. One of these challenges is the so-called "big data skills gap": research projects that use these datasets require a different knowledge base and skill set, and a different set of technical and conceptual tools, than traditional small-scale experimental studies that have been the mainstay of human neuroscience for many years [5]. These skills and tools augment research on the human brain with elements of software engineering, automation, scalable computing, provenance tracking, and advanced statistical and data visualization methods, which are sometimes collectively known as "data science" [6, 7]. While there is a growing appreciation for the importance of data science tools and skills in neuroscience research, there is still a dearth of opportunities for students and researchers in neuroscience to learn about them.

One of the ways that human neuroscience has responded to the need to develop these skills has been by organizing hackathons. These are time-bounded events where participants come together to propose and work on ad-hoc (usually technical) solutions to a range of challenges. In scientific contexts, hackathons have been taken up as a way for technically-oriented researchers to collaborate on software for data analysis, and come up with creative solutions. In particular, the Brainhack organization (https://brainhack.org/) has organized a range of hackathon events around human neuroscience research since 2012, both in tandem with the Organization for Human Brain Mapping (since 2013), and in local and distributed formats as well [8, 9]. In addition to convening technically-oriented researchers, and by co-locating them with major conferences, these events were designed to be welcoming to a broader audience and to provide opportunities for cross-over between technical developments and the broader neuroimaging community.

Meanwhile, another broader response to the need for data science training was being offered by grassroots efforts for peer learning of the various skills. Prominent among these efforts, The Carpentries (originally also known as "Software Carpentry"), is a volunteer organization founded in 1998 that provides researchers with practical training in computing skills, with a focus on automation, data management, and basic programming [10]. It is hard to overestimate the impact that Software Carpentry has had. Over the years it has existed, Software Carpentry has expanded to include a veritable army of more than 1,200 volunteer instructors worldwide, who have taught more than 2,000 workshops in 56 countries since 2012, reaching well over 26,000 learners. In addition, its open-source materials are freely available, contributing to a broader movement promoting computational literacy in the research community.

1.2 A brief history of NeuroHackademy prior to the COVID-19 pandemic

NeuroHackademy is a summer institute in neuroimaging and data science, established in 2016. It was started as a response to the big data skills gap and as a partial solution to

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

the training needs of researchers in human neuroscience. In essence, the course merges the Brainhack approach with the Software Carpentry approach in the same setting: The course takes an innovative approach to training, by splitting the time devoted to the course into two phases. The first phase is designed as an explicit learning phase in which participants attend lectures about the application of data science methods to topics in human neuroscience. These lectures include both discussions of general themes like neuroethics or data governance and tutorials on specific topics, such as introductions to programming, software engineering, data visualization, and machine learning. The second phase of the course is an implicit project-based learning phase. In this phase, participants engage in a hackathon, where they can propose projects and ideas to work on, and form teams together with other participants [11]. During this phase, the course instructors and organizers serve as mentors in an as-needed capacity, giving participants an opportunity to learn by doing. The efficacy of this kind of active learning is supported by extensive research, including by a metaanalysis of 225 studies of the effects of active learning on failure rates in college-level STEM lectures [12]. Similarly, collaboration and problem-based learning are considered effective not only in improving retention, but also in improving student motivation and inclusion of members of under-represented groups [13]. These benefits are particularly salient for the second phase of the event, but the design of the first phase is also motivated by this body of research. Instruction during the first week emphasizes active exploration of the topics, particularly for hands-on tutorials in version control, programming and machine learning. These lessons are inspired and designed based on the teaching principles that guide the Carpentries [14]. In section 2.2 of the paper, we will discuss some of the technical elements that facilitate these hands-on explorations.

NeuroHackademy is hosted by the eScience Institute, an interdisciplinary hub for data science at the University of Washington, and relies on the physical, intellectual, and administrative infrastructure that exists at the Institute. The NeuroHackademy directors (AR and NB) are members of the core research staff at the eScience Institute, and many of the NeuroHackademy instructors are also members of the eScience core research staff, or affiliated with the eScience Institute. The first NeuroHackademy was held in 2016 as a one-week, 40-participant event, held at the University of Washington eScience Institute Data Science Studio. A second similar —40 participant, one-week—format was held in 2017. That year, the project was also awarded an R25 grant from the National Institute of Mental Health, which would support the expansion of the course in 2018 to include more participants (60 participants) and a longer duration (two weeks). This expansion necessitated the migration of the event to another physical location at the University of Washington, the Alder Hall conference center, which includes a large lecture hall, several smaller classrooms and a large modular commons space, which could accommodate a larger and longer event. NeuroHackademy 2019 was held in the same space and in similar format.

2 How do you move a summer school online?

In Spring 2020 there was a lot of uncertainty. That year’s NeuroHackademy event was already intended to take place for the fifth time in July. 70 participants had already been selected from the more than 300 applications to participate in the event, and instructors were already planning what they would teach. But as the COVID-19 pandemic was spreading

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

over the entire world, and as most academic researchers found themselves isolated in their homes for longer and longer periods of time, it became apparent that no large gatherings of people would be possible that summer. As organizers of NeuroHackademy, we strongly considered canceling the event altogether with the intention of reconvening in the summer of 2021. However, an alternative possibility also occurred to us: to adapt the set of online technologies that we so frequently employed in service of distributed work on data science projects (e.g., on GitHub; see section 3.1) to the task of holding a globally distributed online NeuroHackademy event instead. Such an event need not be limited in attendance by the constraints of physical space (e.g., in terms of the number of participants), so we decided to open up the registration to the 2020 event to anyone who was interested in participating, and advertised it through social media and several popular neuroscience mailing lists. Within a few weeks, we had more than 1,400 registrants who had filled out a simple online registration form. Of these, more than 500 participants logged in on the first day of the event, and more than 100 remained on the last day of the event, which was shortened to five days. The new experience of delivering a remote course was eye opening. First of all, it was clear that there was substantial interest in such training regardless of its delivery modality (and maybe particularly when everyone was locked-in at home). Second, we learned that many individuals who would be reluctant to make the trip to Seattle even in normal times, due to a variety of constraints (including the financial burden of travel to Seattle, childcare or other care constraints, and the limitations on accessibility and duration of visa applications to enter the US), may very well prefer the online modality. Similar advantages were identified by other globally distributed neuroscience training efforts such as Neuromatch Academy [15–17] [1]. However, the size of the 2020 event precluded an effective hackathon from taking place, despite an attempt to organize one on the last day of the event. The challenge of creating teams and finding ways to collaborate remotely with such a large number of participants was cacophonous. In the subsequent year, 2021, while in-person on-campus work was slowly resuming, we decided to nevertheless hold an online-only event. Based on the experience of 2020, we tailored the parameters of the event more precisely for this format, offering shorter days of lectures and tutorials in the first week (with the schedule running from 9:30am to 1:30pm or 2pm PT each day) and a more relaxed pace of hacking in the second week. We also reduced the cohort size to only 30 participants this time, allowing for a more organic process of team formation to take place.

2.1 Lessons learned: adapting NeuroHackademy to the post-COVID world

Our experience during 2020 and 2021 taught us two things:

1. We can reach a wider and more diverse audience of people with online and hybrid workshops than we can using in-person workshops alone;
2. It is possible (if not easy) to hold an online workshop that includes a meaningful hackathon experience for online participants.

Moving forward, and as we applied for renewed funding from NIH for five more years of grant support, we proposed to hold a “hybrid” event, where some participants would

[1]The similarity in names does sometimes lead to confusion: Every year since, we have received a few applications to participate in Neuromatch Academy whenever we have put out a call for applications to NeuroHackademy

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

attend in person in Seattle, and others would attend through the Zoom tele-conference software from wherever they were. In 2022, even though the US public health emergency declaration was still in force, we elected to hold the first hybrid in-person and online event. In consideration of the public health emergency, we arranged for a range of COVID precautions, including daily testing by participants and instructors, requesting that participants wear face masks when indoors, and using air purifier devices in all of the spaces where the course took place. The public health emergency officially ended on May 11, 2023, and the second hybrid event that took place in August 2023 included fewer COVID precautions, in line with UW policies and also consistent with many other scientific gatherings that took place in 2023 (e.g., the OHBM conference in Montreal).

2.2 The hybrid format

The hybrid events held in 2022 and 2023 serve as the basis for some preliminary reflections on the advantages and challenges of such events (described in this section), as well as for formulating some of the best practices in organizing these events (described in section 3.2).

The operation of a hybrid hackweek requires substantial coordination among participants, instructors, and administrators. In additional to the administrative task of ensuring that all in-person participants and instructors arrive at the required location at the correct time, administrators must ensure that remote participants and instructors are able to join the hackweek's events, that announcements and event details can be found by both types of participants (remote and in-person), that lectures are not only viewable by but also inclusive of both types of participants, that instructors are able to provide lecture materials and tutorials in a format that is accessible to and usable by all participants, that all participants are able to access and use all lecture materials, and that both types of participants are able to participate meaningfully in the hackathon portion of the workshop. We consider these challenges by dividing them into three categories: challenges related to teaching and communication, challenges related to computation and infrastructure, and challenges related to remote/in-person integration.

Teaching and communication.—In early 2020, existing tools for remote teaching and communication proliferated and improved rapidly due to the COVID-19 pandemic, and the fruits of this proliferation provided much of the technological foundation for running a hybrid event. The primary technological tools used for communicating with students in the lead-up to the event and about day-to-day matters during the event was Slack (https://slack.com/). Slack has many advantages, one of which is that it is broadly used in industry for inter- and intra-team communication. We chose to use Slack because it is a versatile communication tool that enables both real-time chat during lectures or hack-work as well as easy asynchronous communication between instructors, administrators, and participants throughout the event. During the event, participants and administrators can create specific channels, which are used to review lectures, to coordinate hackathon teams, and to organize other events such as lightning talks or breakout sessions.

Lectures themselves were held over Zoom (https://zoom.com/): remote lecturers simply logged into a preset Zoom room in which lectures were scheduled and delivered the lecture

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

using screen-sharing for slides; in-person lectures were given from a computer that was also logged into the preset lecture Zoom room. Communication between lecturers and participants could also be facilitated over Zoom, with in-person speakers and microphones specifically wired such that questions from the remote audience that could be heard by in-person participants and vice versa. Although Zoom offers a "webinar" mode specifically for large lectures, we found that the standard meeting mode was sufficient and encouraged conversations during lectures.

Computation and infrastructure.—A substantial challenge directly related to hackathons and technical workshops is ensuring that all participants are able to follow lectures—especially live-coding lectures—and to work together using shared software tools. Traditionally, this task is extremely difficult because variation in the operating systems and software configurations among participants and instructors is very high. An instructor presenting a tutorial on git, for example, would typically need to account for small changes in the lecture due to some participants using Windows or Mac, rather than, for example, Linux. For lessons using bespoke or idiosyncratic scientific libraries, many of which are not designed for a broad range of platforms, these challenges can be substantially magnified.

One solution to the problem of diverse computational environments is the use of a shared server or cloud platform. There are many options for providing students with logins to a shared and managed computational environments, but many of these options create substantial overhead for the hackweek administrators. For example, procuring and administering a private server is not only expensive with respect to monetary cost but also highly demanding in terms of administrative cost. Such a server would also be fundamentally inelastic and thus would likely be underused during lectures but heavily used during hack projects.

Fortunately, contemporary cloud computing provides a similar but substantially more elastic solution in the form of bespoke cloud-based servers [18]. Because Neuro-Hackademy focuses on teaching Python and Python-based libraries, we commissioned a managed JupyterHub server from the International Interactive Computing Consortium (https://2i2c.org/), a non-profit organization devoted to making tools from the Jupyter ecosystem accessible to a wide and international audience. This managed server approach had several substantial advantages. Primarily, it allowed us to provide a standard computation environment, including both the ecosystem of installed software available to students and the datasets (or lecture aids) that were needed for demonstrations or examples. This allowed instructors to prepare live-coding sessions or demonstrations during their lectures without worrying that some subset of students would be unable to follow along. Additionally, because the JupyterHub server operated in the cloud, compute resources and storage space could be dynamically adjusted during the instruction and hackathon portions of the workshop, based on the participants' needs. Because JupyterHub grants both the ability to engage with Python using browser-based notebooks and the ability to interact with the underlying Unix server via browser-based terminals, all hackathon activities, such as coordinating code on GitHub, can be accomplished using the server. Finally, the use of a JupyterHub server for virtually all hackweek activities creates an equity of access to course materials and activities that is otherwise difficult to ensure. Participants in a hackweek,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

whether in-person or remote, are likely to have very different computational resources at their disposal, and it is paramount to organize the course such that participants without a powerful personal computer are not excluded from activities.

Hybrid in-person/remote integration.—A critical task for the administrators of a hybrid hackweek is ensuring that in-person and remote participants are able to coordinate and integrate with each other throughout the experience. To a large extent, this goal can be accomplished using the technology mentioned above, such as Zoom and Slack. In the case of Slack, so long as official communications are made over Slack, all participants should have equal access to all information. Similarly, the use of Zoom as the primary lecture and meeting venue enables all students to participate. However, one reality of hybrid hackweeks is that some students will be attending in the same timezone as the in-person participants while others will be in very remote timezones. It is thus not sufficient to merely hold lectures online. In such a case, recording and promptly posting videos of lectures online helps remote students without the ability to attend lectures synchronously get the most out of the course. Zoom fortunately has the ability to automatically record meetings to the cloud, and such recordings can be posted to the eScience Institute's YouTube channel (https://www.youtube.com/@UWeScienceInstitute) usually on the same day.

Enabling seamless communication between participants on hybrid hackathon teams is a separate challenge from coordinating lectures. During the hackathon portion of NeuroHackademy, we provided hybrid teams with OWLs (https://owllabs.com/). OWLs are remote communication devices intended for hybrid round-table meetings that serve as both a camera/microphone and a speaker. Using an OWL, the in-person portion of a team can sit at a table and converse easily with team members who are online on Zoom. OWLs also make excellent cameras and speakers for hybrid lectures, in which they serve as the primary conduit between the in-person students and the online students, who can ask questions through the OWL's speaker/microphone interface.

3 Discussion and conclusion

3.1 The importance of a community of practice

Although an event like NeuroHackademy covers some of the needs in terms of training and skills that were mentioned in section 1.1, it also serves a host of other needs. In particular, data science skills are tightly bound with the notion of a *community of practice* (CoP), first formalized by cognitive theorists Lave and Wenger [19]. According to their definition, CoPs are groups of people who share an interest and learn from each other through continuous interaction: a fluid weaving-together of learning, working and innovating [20]. In this theory, learning certain knowledge and skills occurs by becoming a practitioner within a field. Embodying this idea, NeuroHackademy aims to connect participants in the event to a CoP of likely collaborators and colleagues. This notion is specifically supported by the burgeoning existing CoP around the use of data science in neuroimaging, which has a long history stemming from the data-intensive nature of the field [5, 21].

CoPs have also received a significant boost from the adoption of the specific practices common in many open-source software projects related to data science more generally.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Many of the tools for managing and analyzing large datasets are created and maintained by distributed networks of dozens or hundreds of individuals contributing to publicly maintained open-source codebases. As a paradigmatic example, the Scipy package for Scientific computing [22] is very widely used (the paper describing the software has more than 23,000 citations since its publcation in 2020, according to Google Scholar). It has been in continuous development since 2021 and currently counts more than 1,400 unique contributors. The development of the software is still very active with dozens of new commits to the codebase every week. An annual conference (https://conference.scipy.org/) serves as a focal point for this community of developers and an extended community of users. Thanks to their collaborative nature and the fact that they are open to participation, CoPs that transcend geographic and institutional boundaries, such as the Scipy community, have driven important recent innovations in data science. One of the main tools used by these communities is the GitHub online platform. This platform facilitates asynchronous, open and collaborative development of the software tools that are foundational to data science. The distributed and asynchronous nature of their development is among the causes for democratization of these tools. One particular bridge-head for the adoption and acculturation of these practices in human neuroscience occurred surrounding the development of tools for neuroimaging in the Python programming language [23], building upon the foundation laid by the Scipy community.

By contrast, the vast majority of neuroimaging results and analysis tools are generated by a very small number of people working in private. An explicit goal of NeuroHackademy, which is emphasized by the hands-on second week of team projects, is to change this culture of neuroscience by introducing neuroimaging researchers to the tools and practices that allow projects in data science to proceed in an extremely rapid and highly collaborative manner. Recent research on CoPs demonstrates that a CoP grows organically and thrives through the process of “thinking together,” in which practitioners share their understandings of problems in areas of mutual interest with each other and offer guidance. These exchanges do not just demonstrate skills; they also indirectly teach novices the conventions and idiosyncrasies of a practice that are difficult to communicate through formal instruction [24]. By using hands-on tutorials and collaborative hackweek projects that guide the participants of NeuroHackademy through “legitimate peripheral participation” in data science—i.e., small meaningful activities of professional data science practice—it can lower the barrier-to-entry into the CoPs that exist at the intersection of neuroimaging and data science. Notably, many NeuroHackademy hackathon projects have involved legitimate peripheral participation through contributions to existing projects in which course faculty were deeply ingrained. For example, through engagement with the Brain Imaging Data Structure (BIDS) [25, 26], a CoP focused on standards for organizing and sharing neuroimaging data, which comprises data scientists and neuroscientists from all over the world (including several course faculty). NeuroHackademy also aims to empower participants to pass on the habits and knowledge they have acquired at NeuroHackademy to their colleagues and trainees when they return to their home institutions through the CoP approach.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

3.2 Dos and don’ts of a hybrid event

We believe that NeuroHackademy has been a resounding success in each of the eight years that it has been held, including, in particular, the two recent years in which it was held as a hybrid event. This success has been due in no small part to the enthusiasm that the students and instructors bring to the event and has occurred despite many mistakes, misconceptions, and lessons that we as administrators have learned during its execution. Here, we attempt to distill several of these lessons into high-level advice for those who wish to translate the success we have observed to other educational events. We focus here on advice specifically about hybrid event administration.

Technology.—The role of technology in a hybrid event should be to facilitate and simplify hybrid activities such as tutorials and teamwork. The pandemic greatly accelerated the availability and adoption of such technology, such that students need not arrive with any software installed aside from a web-browser. With reasonable planning, the technological challenges stereotypically faced by participants in workshops and classrooms can be nearly eliminated from such an event.

DO provide a common computational environment (coordinating lessons and tutorials across disparate environments and operating systems is very difficult).

DO Use technology that facilitates group interactions over teleconference. For example, NeuroHackademy relies heavily on OWL cameras that provide a wide field of view and better sound over teleconference.

DO make sure official communications happen in a hybrid-friendly space (like Slack).

DO think about what participants will be able to do after the event is over. For example, how to translate software and data from a Jupyterhub to their own computers.

Hybrid-friendly activities.—Certain scientific programming such as poster sessions are common at educational events like NeuroHackademy, but not all such events are equally useful for an in-person and a hybrid cohort. When choosing what additional activities to include in a hybrid program, it is critical that administrators consider the experience from the perspective of both online and in-person participants.

DON’T organize a poster session. Though this is an opportunity to increase visibility of all participants [27], we have found that this tends to split the group into two, leaving many of the online participants feeling particularly excluded.

DO have lightning talks instead of a poster session. This also provides an opportunity for all participants to be visible, but gives a much more similar experience for remote and in-person participants [2].

[2]Including an opportunity to present their own work is also helpful to some students, because many departments provide travel support for students on the condition that they present at a meeting/event

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Anticipating the differences.—Many of the difficulties inherent in participating remotely in a large educational event arise from mismatches between the expectations of participants and those of the administrators. In many cases, these expectations can be mitigated with minimal planning.

DON'T pretend that online and in-person participants have the same experience. Both may have a valuable experience, but they are not identical.

DON'T require remote participants to attend events synchronously, allowing them to operate in their own time zone. For example, uploading videos quickly to be viewed asynchronously is crucial.

DO Include several participants in each remote time-zone that you have, so that they have a cohort that they can interact with synchronously, even during times that activities are not taking place in-person at the main event site, because of the late hour. This is akin to Neuromatch Academy's "pod" structure [17], which forms and matches groups of participants in each time-zone based on their interests.

DON'T require all instructors to attend in-person; make sure that instructors have an online presence (e.g., through Zoom office hours).

Creating a community.—Finally, one of the greatest challenges of hybrid education events is ensuring that online and in-person participants have similar access to the community that inherently grows out of participant interactions during group learning and projects. While it is unlikely that in-person and remote participants will have similar exposure to such communities through the event, some strategies can explicitly encourage such access.

DO build in ways for online and in-person participants to interact with each other (e.g., in hybrid ice-breaker sessions).

DO ensure that online participants can be involved in hack-groups from the beginning; i.e., make sure that initial group formation happens in a hybrid-friendly space rather than using an inherently in-person medium such as a whiteboard.

DO take advantage of existing communities to foster legitimate peripheral participation. For example, the Neurostars website (https://neurostars.org) is a Q&A website for the global neuroinformatics community, where discussions that start at NeuroHackademy could continue and provide a place for participants to continue their involvement in the CoP around data science and neuroimaging.

DO Make an effort to create "online interaction rituals", This concept refers to repeated activities that happen online and form a shared experience [28]. For example, in one of the first NeuroHackademy events, one of the instructors would kick off every day in the event by posting a picture of a squirrel to Slack. At the hybrid events, we formalize an online interaction ritual by starting each day with a "stand-up" session, where each project team reports on the progress of their work.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Acknowledgments.

The authors wish to acknowledge Tal Yarkoni's important role in establishing the course, and the contributions of many instructors who have devoted significant effort and time to the success of the event. We are also grateful to NeuroHackademy participants for their engagement and for all the feedback they have provided over the years.

Funding Declaration.

NeuroHackademy is funded through grant R25MH112480 from the National Institue of Mental Health. The two first installations of the course were additionally funded by a grant from the Gordon & Betty Moore Foundation and the Alfred P. Sloan Foundation to the University of Washington eScience Institute Data Science Environment.

References

[1]. Van Essen DC, Ugurbil K, Auerbach E, Barch D, Behrens TEJ, Bucholz R, Chang A, Chen L, Corbetta M, Curtiss SW, Penna SD, Feinberg D, Glasser MF, Harel N, Heath AC, Larson-Prior L, Marcus D, Michalareas G, Moeller S, Oostenveld R, Petersen SE, Prior F, Schlaggar BL, Smith SM, Snyder AZ, Xu J, Yacoub E: "the human connectome project: A data acquisition perspective ". Neuroimage 62(4), 2222–2231 (2012) 10.1016/j.neuroimage.2012.02.018 [PubMed: 22366334]

[2]. Glasser MF, Smith SM, Marcus DS, Andersson JLR, Auerbach EJ, Behrens TEJ, Coalson TS, Harms MP, Jenkinson M, Moeller S, Robinson EC, Sotiropoulos SN, Xu J, Yacoub E, Ugurbil K, Van Essen DC: The human connectome project's neuroimaging approach. Nat. Neurosci 19(9), 1175–1187 (2016) [PubMed: 27571196]

[3]. Jernigan TL, Brown SA, Dowling GJ: The adolescent brain cognitive development study. J. Res. Adolesc 28(1), 154–156 (2018) [PubMed: 29460352]

[4]. Mennes M, Biswal BB, Castellanos FX, Milham MP: Making data sharing work: the FCP/INDI experience. Neuroimage 82, 683–691 (2013) [PubMed: 23123682]

[5]. Van Horn JD: Opinion: Big data biomedicine offers big higher education opportunities. Proc. Natl. Acad. Sci. U. S. A 113(23), 6322–6324 (2016) [PubMed: 27274038]

[6]. Donoho D: 50 years of data science. J. Comput. Graph. Stat 26(4), 745–766 (2017)

[7]. Cleveland WS: Data science: an action plan for expanding the technical areas of the field of statistics. Int. Stat. Rev 69(1), 21–26 (2001)

[8]. Cameron Craddock R, S Margulies D, Bellec P, Nolan Nichols B, Alcauter S, A Barrios F, Burnod Y, J Cannistraci C, Cohen-Adad J, De Leener B, Dery S, Downar J, Dunlop K, R Franco A, Seligman Froehlich C, J Gerber A, S Ghosh S, J Grabowski T, Hill S, Sólon Heinsfeld A, Matthew Hutchison R, Kundu P, R Laird A, Liew S-L, J Lurie D, G McLaren D, Meneguzzi F, Mennes M, Mesmoudi S, O'Connor D, H Pasaye E, Peltier S, Poline J-B, Prasad G, Fraga Pereira R, Quirion P-O, Rokem A, S Saad Z, Shi Y, C Strother S, Toro R, Q Uddin L, D Van Horn J, W Van Meter J, C Welsh R, Xu T: Brainhack: a collaborative workshop for the open neuroscience community. Gigascience 5, 16 (2016) [PubMed: 27042293]

[9]. Gau R, Noble S, Heuer K, Bottenhorn KL, Bilgin IP, Yang Y-F, Huntenburg JM, Bayer JMM, Bethlehem RAI, Rhoads SA, Vogelbacher C, Borghesani V, Levitis E, Wang H-T, Van Den Bossche S, Kobeleva X, Legarreta JH, Guay S, Atay SM, Varoquaux GP, Huijser DC, Sandström MS, Herholz P, Nastase SA, Badhwar A, Dumas G, Schwab S, Moia S, Dayan M, Bassil Y, Brooks PP, Mancini M, Shine JM, O'Connor D, Xie X, Poggiali D, Friedrich P, Heinsfeld AS, Riedl L, Toro R, Caballero-Gaudes C, Eklund A, Garner KG, Nolan CR, Demeter DV, Barrios FA, Merchant JS, McDevitt EA, Oostenveld R, Craddock RC, Rokem A, Doyle A, Ghosh SS, Nikolaidis A, Stanley OW, Uruñuela E, Anousheh N, Arnatkeviciute A, Auzias G, Bachar D, Bannier E, Basanisi R, Basavaraj A, Bedini M, Bellec P, Benn RA, Berluti K, Bollmann S, Bollmann S, Bradley C, Brown J, Buchweitz A, Callahan P, Chan MY, Chandio BQ, Cheng T, Chopra S, Chung AW, Close TG, Combrisson E, Cona G, Constable RT, Cury C, Dadi K, Damasceno PF, Das S, De Vico Fallani F, DeStasio K, Dickie EW, Dorfschmidt L, Duff EP, DuPre E, Dziura S, Esper NB, Esteban O, Fadnavis S, Flandin G, Flannery JE, Flournoy J, Forkel SJ, Franco AR, Ganesan S, Gao S, García Alanis JC, Garyfallidis E, Glatard T, Glerean E, Gonzalez-Castillo J, Praag CD, Greene AS, Gupta G, Hahn CA, Halchenko YO, Handwerker D,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Hartmann TS, Hayot-Sasson V, Heunis S, Hoffstaedter F, Hohmann DM, Horien C, Ioanas H-I, Iordan A, Jiang C, Joseph M, Kai J, Karakuzu A, Kennedy DN, Keshavan A, Khan AR, Kiar G, Klink PC, Koppelmans V, Koudoro S, Laird AR, Langs G, Laws M, Licandro R, Liew S-L, Lipic T, Litinas K, Lurie DJ, Lussier D, Madan CR, Mais L-T, Mansour L S, Manzano-Patron JP, Maoutsa D, Marcon M, Margulies DS, Marinato G, Marinazzo D, Markiewicz CJ, Maumet C, Meneguzzi F, Meunier D, Milham MP, Mills KL, Momi D, Moreau CA, Motala A, Moxon-Emre I, Nichols TE, Nielson DM, Nilsonne G, Novello L, O'Brien C, Olafson E, Oliver LD, Onofrey JA, Orchard ER, Oudyk K, Park PJ, Parsapoor M, Pasquini L, Peltier S, Pernet CR, Pienaar R, Pinheiro-Chagas P, Poline J-B, Qiu A, Quendera T, Rice LC, Rocha-Hidalgo J, Rutherford S, Scharinger M, Scheinost D, Shariq D, Shaw TB, Siless V, Simmonite M, Sirmpilatze N, Spence H, Sprenger J, Stajduhar A, Szinte M, Takerkart S, Tam A, Tejavibulya L, Schotten M, Thome I, Silva L, Traut N, Uddin LQ, Vallesi A, VanMeter JW, Vijayakumar N, Oleggio Castello MV, Vohryzek J, Vukojević J, Whitaker KJ, Whitmore L, Wideman S, Witt ST, Xie H, Xu T, Yan C-G, Yeh F-C, Yeo BTT, Zuo X-N: Brainhack: Developing a culture of open, inclusive, community-driven neuroscience. Neuron (2021)

[10]. Wilson G: Software carpentry: lessons learned. F1000Res. 3 (2014)

[11]. Huppenkothen D, Arendt A, Hogg DW, Ram K, VanderPlas JT, Rokem A: Hack weeks as a model for data science education and collaboration. Proc. Natl. Acad. Sci. U. S. A 115(36), 8872–8877 (2018) [PubMed: 30127025]

[12]. Freeman S, Eddy SL, McDonough M, Smith MK, Okoroafor N, Jordt H, Wenderoth MP: Active learning increases student performance in science, engineering, and mathematics. Proc. Natl. Acad. Sci. U. S. A 111(23), 8410–8415 (2014) [PubMed: 24821756]

[13]. Prince M: Does active learning work? a review of the research. J. Eng. Educ 93(3), 223–231 (2004)

[14]. Wilson G: Teaching Tech Together: How to Make Your Lessons Work and Build a Teaching Community Around Them. CRC Press, ??? (2019)

[15]. Kording KP: For love of neuroscience: The neuromatch movement. Neuron 109(19), 3034–3035 (2021) [PubMed: 34559980]

[16]. Viegen T, Akrami A, Bonnen K, DeWitt E, Hyafil A, Ledmyr H, Lindsay GW, Mineault P, Murray JD, Pitkow X, Puce A, Sedigh-Sarvestani M, Stringer C, Achakulvisut T, Alikarami E, Atay MS, Batty E, Erlich JC, Galbraith BV, Guo Y, Juavinett AL, Krause MR, Li S, Pachitariu M, Straley E, Valeriani D, Vaughan E, Vaziri-Pashkam M, Waskom ML, Blohm G, Kording K, Schrater P, Wyble B, Escola S, Peters MAK: Neuromatch academy: Teaching computational neuroscience with global accessibility. Trends Cogn. Sci 25(7), 535–538 (2021) [PubMed: 33994097]

[17]. Hart B, Achakulvisut T, Adeyemi A, Akrami A, Alicea B, Alonso-Andres A, Alzate-Correa D, Ash A, Ballesteros J, Balwani A, Batty E, Beierholm U, Benjamin A, Bhalla U, Blohm G, Blohm J, Bonnen K, Brigham M, Brunton B, Butler J, Caie B, Gajic N, Chatterjee S, Chavlis S, Chen R, Cheng Y, Chow H.m., Chua R, Dai Y, David I, DeWitt E, Denis J, Dipani A, Dorschel A, Drugowitsch J, Dwivedi K, Escola S, Fan H, Farhoodi R, Fei Y, Fiquet P-É, Fontolan L, Forest J, Fujishima Y, Galbraith B, Galdamez M, Gao R, Gjorgjieva J, Gonzalez A, Gu Q, Guo Y, Guo Z, Gupta P, Gurbuz B, Haimerl C, Harrod J, Hyafil A, Irani M, Jacobson D, Johnson M, Jones I, Karni G, Kass R, Kim H, Kist A, Koene R, Kording K, Krause M, Kumar A, Kühn N, Lc R, Laporte M, Lee J, Li S, Lin S, Lin Y, Liu S, Liu T, Livezey J, Lu L, Macke J, Mahaffy K, Martins A, Martorell N, Martínez M, Mattar M, Menendez J, Miller K, Mineault P, Mohammadi N, Mohsenzadeh Y, Morgenroth E, Morshedzadeh T, Mosberger A, Muliya M, Mur M, Murray J, Nd Y, Naud R, Nayak P, Oak A, Castillo I, Orouji S, Otero-Millan J, Pachitariu M, Pandey B, Paredes R, Parent J, Park I, Peters M, Pitkow X, Poirazi P, Popal H, Prabhakaran S, Qiu T, Ragunathan S, Rodriguez-Cruces R, Rolnick D, Sahoo A, Salehinajafabadi S, Savin C, Saxena S, Schrater P, Schroeder K, Schwarze A, Sedigh-Sarvestani M, Sekhar K, Shadmehr R, Shanechi M, Sharma S, Shea-Brown E, Shenoy K, Shimabukuro C, Shuvaev S, Sin M, Smith M, Steinmetz N, Stosio K, Straley E, Strandquist G, Stringer C, Tomar R, Tran N, Triantafillou S, Udeigwe L, Valeriani D, Valton V, Vaziri-Pashkam M, Vincent P, Vishne G, Wallisch P, Wang P, Ward C, Waskom M, Wei K, Wu A, Wu Z, Wyble B, Zhang L, Zysman D, Uquillas F, Viegen T: Neuromatch academy: a 3-week, online summer school in computational neuroscience. J. Open Source Educ 5(49), 118 (2022)

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

[18]. Holdgraf C, Culich A, Rokem A, Deniz F, Alegro M, Ushizima D: Portable learning environments for hands-on computational instruction: Using container- and cloud-based technology to teach data science. In: Proceedings of the Practice and Experience in Advanced Research Computing 2017 on Sustainability, Success and Impact, pp. 1–9 (2017)

[19]. Lave J, Wenger E: Situated Learning: Legitimate Peripheral Participation vol. 138. Cambridge University Press, New York, NY, US (1991)

[20]. Brown JS, Duguid P: Organizational learning and Communities-of-Practice: Toward a unified view of working, learning, and innovation. Organization Science 2(1), 40–57 (1991)

[21]. Van Horn JD: Bridging the brain and data sciences. Big Data (2020)

[22]. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, Walt SJ, Brett M, Wilson J, Millman KJ, Mayorov N, Nelson ARJ, Jones E, Kern R, Larson E, Carey CJ, Polat İ, Feng Y, Moore EW, VanderPlas J, Laxalde D, Perktold J, Cimrman R, Henriksen I, Quintero EA, Harris CR, Archibald AM, Ribeiro AH, Pedregosa F, Mulbregt P, Vijaykumar A, Bardelli AP, Rothberg A, Hilboll A, Kloeckner A, Scopatz A, Lee A, Rokem A, Woods CN, Fulton C, Masson C, Häggström C, Fitzgerald C, Nicholson DA, Hagen DR, Pasechnik DV, Olivetti E, Martin E, Wieser E, Silva F, Lenders F, Wilhelm F, Young G, Price GA, Ingold G-L, Allen GE, Lee GR, Audren H, Probst I, Dietrich JP, Silterra J, Webber JT, Slavič J, Nothman J, Buchner J, Kulick J, Schönberger JL, Miranda Cardoso J, Reimer J, Harrington J, Rodríguez JLC, Nunez-Iglesias J, Kuczynski J, Tritz K, Thoma M, Newville M, Kümmerer M, Bolingbroke M, Tartre M, Pak M, Smith NJ, Nowaczyk N, Shebanov N, Pavlyk O, Brodtkorb PA, Lee P, McGibbon RT, Feldbauer R, Lewis S, Tygier S, Sievert S, Vigna S, Peterson S, More S, Pudlik T, Oshima T, Pingel TJ, Robitaille TP, Spura T, Jones TR, Cera T, Leslie T, Zito T, Krauss T, Upadhyay U, Halchenko YO, Vázquez-Baeza Y: Scipy 1.0: fundamental algorithms for scientific computing in python. Nature Methods 17(3), 261–272 (2020) 10.1038/s41592-019-0686-2 [PubMed: 32015543]

[23]. Millman KJ, Brett M: Analysis of functional magnetic resonance imaging in python. Computing in Science Engineering 9(3), 52–55 (2007)

[24]. Pyrko I, Dörfler V, Eden C: Thinking together: What makes communities of practice work? Hum. Relat 70(4), 389–409 (2017) [PubMed: 28232754]

[25]. Gorgolewski KJ, Auer T, Calhoun VD, Craddock RC, Das S, Duff EP, Flandin G, Ghosh SS, Glatard T, Halchenko YO, Handwerker DA, Hanke M, Keator D, Li X, Michael Z, Maumet C, Nichols BN, Nichols TE, Pellman J, Poline J-B, Rokem A, Schaefer G, Sochat V, Triplett W, Turner JA, Varoquaux G, Poldrack RA: The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Sci Data 3, 160044 (2016) [PubMed: 27326542]

[26]. Poldrack RA, Markiewicz CJ, Appelhoff S, Ashar YK, Auer T, Baillet S, Bansal S, Beltrachini L, Benar CG, Bertazzoli G, Bhogawar S, Blair RW, Bortoletto M, Boudreau M, Brooks TL, Calhoun VD, Castelli FM, Clement P, Cohen AL, Cohen-Adad J, D'Ambrosio S, Hollander G, Iglesia-Vayá M, Vega A, Delorme A, Devinsky O, Draschkow D, Duff EP, DuPre E, Earl E, Esteban O, Feingold FW, Flandin G, Galassi A, Gallitto G, Ganz M, Gau R, Gholam J, Ghosh SS, Giacomel A, Gillman AG, Gleeson P, Gramfort A, Guay S, Guidali G, Halchenko YO, Handwerker DA, Hardcastle N, Herholz P, Hermes D, Honey CJ, Innis RB, Ioanas H-I, Jahn A, Karakuzu A, Keator DB, Kiar G, Kincses B, Laird AR, Lau JC, Lazari A, Legarreta JH, Li A, Li X, Love BC, Lu H, Maumet C, Mazzamuto G, Meisler SL, Mikkelsen M, Mutsaerts H, Nichols TE, Nikolaidis A, Nilsonne G, Niso G, Norgaard M, Okell TW, Oostenveld R, Ort E, Park PJ, Pawlik M, Pernet CR, Pestilli F, Petr J, Phillips C, Poline J-B, Pollonini L, Raamana PR, Ritter P, Rizzo G, Robbins KA, Rockhill AP, Rogers C, Rokem A, Rorden C, Routier A, Saborit-Torres JM, Salo T, Schirner M, Smith RE, Spisak T, Sprenger J, Swann NC, Szinte M, Takerkart S, Thirion B, Thomas AG, Torabian S, Varoquaux G, Voytek B, Welzel J, Wilson M, Yarkoni T, Gorgolewski KJ: The past, present, and future of the brain imaging data structure (BIDS). ArXiv (2023)

[27]. Levitis E, Praag CD, Gau R, Heunis S, DuPre E, Kiar G, Bottenhorn KL, Glatard T, Nikolaidis A, Whitaker KJ, al., E.: Centering inclusivity in the design of online conferences (2021)

[28]. Simons I: Events and online interaction: the construction of hybrid event communities. Leisure Studies 38(2), 145–159 (2019)