

1 *Stefano Moia^{1, 2, 3}, Hao-Ting Wang^{1, 4}, Anibal S. Heinsfeld^{5, 6}, Dorota Jarecka⁷, Yu-Fang Yang⁸, Stephan Heunis⁹,*
2 *Michele Svanera¹⁰, Benjamin De Leneer^{1, 11, 12, 13}, Andrea Gondova^{1, 14, 15}, Sin Kim¹⁶, Arshitha Basavaraj¹⁷, Jo-*
3 *hanna M. Bayer^{1, 18, 19}, Roza G. Bayrak^{1, 20}, Pierre-Louis Bazin^{21, 22}, Isil Poyraz Bilgin^{4, 23}, Steffen Bollmann^{24, 25},*
4 *Daniel Borek²⁶, Valentina Borghesani^{1, 27}, Trang Cao²⁸, Gang Chen²⁹, Alejandro De La Vega³⁰, Sebastian Dresbach³¹,*
5 *Philipp Ehses³², Jan Ernsting^{33, 34, 35}, Inês Esteves³⁶, Oscar Ferrante³⁷, Kelly G. Garner³⁸, Rémi Gau³⁹, Elodie*
6 *Germani⁴⁰, Tara Ghafari³⁷, Satrajit S. Ghosh^{7, 41}, Sarah Elizabeth Goodale⁴², Cassandra D. Gould van Praag^{43, 44},*
7 *Samuel Guay⁴⁵, Omer Faruk Gulban^{46, 31}, Yaroslav O. Halchenko⁴⁷, Michael Hanke^{9, 48}, Peer Herholz⁴⁹, Katja Heuer⁵⁰,*
8 *Felix Hoffstaedter^{9, 51}, Ruoqi Huang⁵², Renzo Huber⁵³, Ole Jensen³⁷, Kan Keeratimahat⁵⁴, Julian Q. Kosciessa⁵⁵,*
9 *Sladjana Lukic⁵⁶, Neville Magielse^{9, 57, 58}, Christopher J. Markiewicz⁵⁹, Caroline G. Martin⁶⁰, Camille Maumet⁶¹,*
10 *Anna Menacher⁶², Jeff Mentch^{7, 63}, Christian Mönch⁹, Shammi More^{9, 48}, Leonardo Muller-Rodriguez^{64, 65}, Samuel*
11 *A. Nastase⁶⁶, Eliana Nicolaisen-Sobesky⁹, Dylan M. Nielson⁶⁷, Christopher R. Nolan³⁸, François Paugam^{68, 69}, Pedro*
12 *Pinheiro-Chagas⁷⁰, Ana Luísa Pinho^{71, 72}, Alessandra Pizzuti^{53, 46}, Benjamin Poldrack⁶⁴, Benedikt A. Poser⁵³, Roberta*
13 *Rocca⁷³, Jacob Sanz-Robinson⁷⁴, Kelvin Sarink³⁴, Kevin R. Sitek⁷⁵, Nadine Spychal⁷⁶, Rüdiger Stirnberg³², Michał*
14 *Szczepanik⁹, Mohammad Torabi⁷⁷, Roberto Roberto Toro⁵⁰, Sebastian GW Urchs⁷⁸, Sofie L. Valk^{22, 9, 81}, Adina S.*
15 *Wagner^{9, 80}, Laura K. Waite⁹, Alexander Q. Waite⁹, Lea Waller⁸¹, Tyler James Wishard^{82, 83}, Jianxiao Wu^{9, 48},*
16 *Yuchen Zhou⁸⁴, Janine D. Bijsterbosch^{1, 85}, The physiopy community⁸⁶,*
17
18 1. OHBM Open Science Special Interest Group (SIG); 2. Neuro-X Institute, École polytechnique fédérale de Lausanne,
19 Geneva, Switzerland; 3. Department of Radiology and Medical Informatics (DRIM), Faculty of Medicine, University of
20 Geneva, Geneva, Switzerland; 3. Centre de recherche de l'Institut universitaire de gériatrie de Montréal, Montréal, QC,
21 Canada; 4. Centre de recherche de l'Institut universitaire de gériatrie de Montréal, Montreal, Quebec, Canada; 5. Depart-
22 ment of Psychology, Center for Perceptual Systems, The University of Texas at Austin, Austin, TX, USA; 6. Department
23 of Computer Science, The University of Texas at Austin, Austin, TX, USA; 7. McGovern Institute for Brain Research,
24 Massachusetts Institute of Technology, Cambridge, MA, USA; 8. Division of Experimental Psychology and Neuropsychol-
25 ogy, Department of Education and Psychology, Freie Universität Berlin, Berlin, Germany; 9. Institute of Neuroscience
26 and Medicine, Research Centre Jülich, Jülich, Germany; 10. Centre for Cognitive Neuroimaging, School of Psychology &
27 Neuroscience University of Glasgow, UK; 11. Institute of Biomedical Engineering, Polytechnique Montreal, Montreal, QC,
28 CA; 12. CHU Sainte-Justine Research Center, University of Montreal, Montreal, QC, CA; 13. Department of Computer
29 Engineering and Software Engineering, Polytechnique Montreal, Montreal, CA; 14. UNIACT, NeuroSpin, CEA, Université

30 Paris-Saclay, Paris, France; 15. NeuroDiderot, Inserm, Université Paris Cité, Paris, France; 16. Brain and Cognitive Engin-
31 eering, Korea Advanced Institute of Science and Technology, Daejeon, Korea; 17. Data Science and Sharing Team (DSST),
32 Intramural Research Program, NIMH, Bethesda, USA; 18. Centre for Youth Mental Health, the University of Melbourne,
33 Australia; 19. Orygen, Melbourne, Australia; 20. Department of Computer Science, Vanderbilt University, Nashville TN,
34 USA; 21. University of Amsterdam, Amsterdam, Netherlands; 22. Max Planck Institute for Human Cognitive and Brain
35 Science, Leipzig, Germany; 23. Machine Learning Department and the Neuroscience Institute, Carnegie Mellon University,
36 Pittsburgh, US; 24. School of Information Technology and Electrical Engineering, The University of Queensland, Brisbane,
37 Australia; 25. Centre for Innovation in Biomedical Imaging Technology, University of Queensland, Brisbane, Australia; 26.
38 Department of Data Analysis, Faculty of Psychology and Educational Sciences, Ghent University, Ghent, Belgium; 27.
39 Department of Psychology, Faculty of Psychology and Science of Education, University of Geneva, Geneva, Switzerland;
40 28. School of Psychological Sciences, Monash University, Clayton, VIC, Australia; 29. Scientific and Statistical Computing
41 Core, NIMH, NIH, Department of Health and Human Services, USA; 30. Department of Psychology, University of Texas at
42 Austin, Austin, TX, USA; 31. Maastricht University, Maastricht, Netherlands; 32. German Center for Neurodegenerative
43 Diseases (DZNE), Bonn, Germany; 33. Institute for Geoinformatics, University of Münster, Germany; 34. University of
44 Münster, Institute for Translational Psychiatry, Münster, Germany; 35. Faculty of Mathematics and Computer Science,
45 University of Münster, Germany; 36. ISR-Lisboa and Department of Bioengineering, Instituto Superior Técnico – Univer-
46 sidade de Lisboa, Lisbon, Portugal; 37. Centre for Human Brain Health, School of Psychology, University of Birmingham,
47 UK; 38. School of Psychology, The University of New South Wales, NSW, Australia; 39. Institut de recherche en sci-
48 ences psychologique, Université catholique de Louvain. Louvain la neuve, Belgique; 40. Univ Rennes, Inria, CNRS, Inserm,
49 Rennes, France; 41. Department of Otolaryngology - Head and Neck Surgery, Harvard Medical School, Boston, MA, USA;
50 42. Department of Biomedical Engineering, Vanderbilt University; 43. Alan Turing Institute, London, UK; 44. Department
51 of Psychiatry, Oxford University, Oxford, UK; 45. Department of Psychology, University of Montreal, Canada; 46. Brain
52 Innovation, Maastricht, The Netherlands; 47. Department of Psychological and Brain Sciences, Dartmouth college, New
53 Hampshire, USA; 48. Medical Faculty, Heinrich Heine University Düsseldorf; 49. McGill, Montreal Neurological Institute
54 - Hospital, Montreal, Quebec, Canada; 50. Institut Pasteur, Université Paris Cité, Unité de Neuroanatomie Appliquée et
55 Théorique, F-75015 Paris, France; 51. Institute of Systems Neuroscience, Medical Faculty & University Hospital Düsseldorf,
56 Heinrich-Heine- University, Düsseldorf, Germany; 52. University of Toronto, Toronto, Ontario, Canada; 53. Department of
57 Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands; 54.
58 Nuffield Department of Population Health, University of Oxford, Oxford, UK; 55. Donders Institute for Brain, Cognition

59 and Behaviour, Radboud University, Nijmegen, The Netherlands; 56. Ruth S. Ammon College of Education and Health
60 Sciences, Department of Communication Sciences and Disorders, Adelphi University, San Francisco, CA, USA; 57. Otto
61 Hahn Cognitive Neurogenetics Group, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1A,
62 04103, Leipzig, Germany; 58. Institute of Systems Neuroscience, Medical Faculty, Heinrich Heine University Düsseldorf,
63 Moorenstraße 5 40225 Düsseldorf, Germany; 59. Stanford University, CA, USA; 60. Vanderbilt University, Nashville, TN
64 USA; 61. Inria, Univ Rennes, CNRS, Inserm, Rennes, France; 62. Department of Statistics, University of Oxford, UK; 63.
65 Program in Speech and Hearing Bioscience and Technology, Harvard University, Cambridge, MA, USA; 64. Psychoinform-
66 atics Lab, Forschungszentrum Jülich, Jülich, Germany; 65. Cambridge Institute for Music Therapy Research (CIMTR),
67 Anglia Ruskin University, Cambridge, UK; 66. Princeton Neuroscience Institute, Princeton University, Princeton, NJ,
68 USA; 67. Machine Learning Team, National Institute of Mental Health, Bethesda, MD, USA; 68. University of Montreal,
69 Montreal, Quebec, Canada; 69. Mila - Quebec Artificial Intelligence Institute, University of Montreal, Montreal, Quebec,
70 Canada; 70. Department of Neurology, Memory and Aging Center, University of California, San Francisco, San Francisco,
71 CA, USA; 71. Western Institute for Neuroscience, Western University, London, Ontario, Canada; 72. Department of Com-
72 puter Science, Faculty of Science, Western University, London, Ontario, Canada; 73. Department of Culture, Cognition,
73 and Computation, School of Culture and Society, Aarhus University, Aarhus, Denmark; 74. Neuroscience, McGill Uni-
74 versity, Quebec, Canada; 75. Department of Communication Science and Disorders, School of Health and Rehabilitation
75 Sciences, University of Pittsburgh, Pittsburgh, PA, USA; 76. Department of Informatics, University of Sussex, UK; 77.
76 Graduate Program in Biological and Biomedical Engineering, McGill University, Montreal, Canada; 78. NeuroDataScienc
77 - ORIGAMI lab, Montreal Neurological Institute-Hospital, McGill University, Montreal, Canada; 79. Institute of Systems
78 Neuroscience Heinrich Heine University Düsseldorf; 80. Department of Psychology, Heinrich Heine University Düsseldorf,
79 Düsseldorf, Germany ; 81. Charité Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-
80 Universität zu Berlin, Department of Psychiatry and Neurosciences CCM, Berlin, Germany; 82. Department of Psychiatry
81 and Biobehavioral Sciences, University of California Los Angeles, Los Angeles, California, United States; 83. Department of
82 Neurology, Washington University in St. Louis, St. Louis, Missouri, United States; 84. Department of Psychology, Carnegie
83 Mellon University, Pittsburgh, USA; 85. Department of Radiology, Washington University in St Louis, Saint Louis, Mis-
84 souri, USA; 86. physiopy.github.io

85

86

87 Proceedings of the UHBM BrainHack 2022

Abstract

90 OHBM Brainhack 2022 took place in June 2022. The first hybrid OHBM hack-
91 athon, it had an in-person component taking place in Glasgow and three hubs around
92 the globe to improve inclusivity and fit as many timezones as possible. In the buzz-
93 ing setting of the Queen Margaret Union and of the virtual platform, 23 projects
94 were presented after development. Following are the reports of 14 of those, as well
95 as a recapitulation of the organisation of the event.

96 Introduction

The Organisation of Human Brain Mapping BrainHack (shortened to OHBM BrainHack [herein](#)) is a yearly satellite event of the main OHBM meeting, organised by the Open Science Special Interest Group following the model of Brainhack hackathons¹. Where other hackathons set up a competitive environment based on outperforming other participants' projects, Brainhacks foster a collaborative environment in which participants can freely collaborate and exchange ideas within and between projects.

This edition of the OHBM Brainhack, that ran across the world over four days, was particularly special for two reasons: it celebrated the tenth year anniversary of Brainhack, and, like the main OHBM conference, it was the first edition to feature an in-person event after two years of virtual events. For this reasons, the whole organisation rotated around five main principles:

- 108 1. Providing a hybrid event incorporating the positive aspects of in-person and virtual
109 events alike,

110 2. Celebrating the 10th anniversary of the Brainhack by bringing back newcomer-
111 friendly hands-on hacking and learning experience, enhancing the Hacktrack and
112 formatting the Traintrack as a collection of materials to consult beforehand and as
113 spontaneous meetings of the participants aimed to learn together.

- 114 3. Bridging the gap between the Brainhack community and the main neuroimaging
115 software developer groups, e.g. AFNI, FSL, SPM,
- 116 4. Due to amount of work required to meet the previous three principles, incorporat-
117 ing from the beginning a team of core organisers with a democratic approach to
118 organisation, with a member in charge of an aspect of the event,
- 119 5. Brainhack event organisation should always be experimental, trying different solu-
120 tions and formats to find a way to improve Brainhack events overall.

121 After a quick explanation of each main contribution of the core team, the next pages
122 are dedicated to the summaries of the projects that were developed during the four days
123 of hacking.

124 1 Hacktrack

125 *Dorota Jarecka, Yu-Fang Yang, Hao-Ting Wang, Stefano Moia*

126

127 The key component of each Brainhack is hacking. The hacking part, known as hack-
128 track, is where attendees collaborate on projects and explore their own ideas. There are
129 4 elements of hacktrack that were organised: project submission, project pitch, hacking
130 period and project summary. For the project submission, we used the GitHub issue sub-
131 mission process that was used during recent years. We updated and simplified a project
132 template from previous years and asked project leaders to open an issue for each project.
133 Each issue after quick check was approved by the moderators and automatic workflows
134 written by the team were responsible for sending project descriptions to the Brainhack
135 page and setting Discord's channels. We received 38 projects that were submitted using
136 this system. The project pitch was set for the morning of the first day and everyone had 2
137 minutes to talk about the suggested project and possible collaborations. After the pitches
138 people had a chance to talk to each other and join the projects they were interested in.
139 This year, we tried to maximise the time for hacking by providing a sparse schedule for

140 talks. The closing ceremony of the Brainhack featured 23 project reports, in which teams
141 talked about their experiences and described the work they accomplished.

142 This edition we allowed remote attendance from other locations. We organised three
143 hubs aiming to cover all time zones, including 1) Asia-Pacific, 2) Glasgow, Europe, Middle
144 East, and Africa, and 3) the Americas, to foster inclusiveness in the hybrid conference
145 format. We also ensured that each hub had one live streamed session with the physical
146 hub in Glasgow.

147 2 Traintrack

148 *Yu-Fang Yang, Dorota Jarecka*

149

150 Traintrack is the educational component of Brainhack events. The aim is to introduce
151 tools and skills for attendees to start hacking. Unlike conventional scientific educational
152 workshops centred around lectures and talks, data science skills are better learned through
153 hands-on experience than lectures. With the Brainhack community growing mature, the
154 community has developed their own curated educational material. *Brainhack School* has
155 supplied high-quality content for independent study on a variety of themes.

156 This year, we combine the collaborative nature of brainhack projects and educational
157 content to reimagine the format of taintack. Thus, we replaced tutorial lectures in the
158 previous editions with curated online educational contents, released them prior to the
159 main event, and attempted to integrate them with the hacktrack projects. This format
160 also provides more time (i.e. schedule) and space (i.e. minimising large space not used for
161 hacking) for attendees to self-organise. Participants were encouraged to form study groups
162 on five suggested topics: 1) setting up your system for analysis 2) python for data analysis,
163 3) machine learning for neuroimaging, 4) version control systems, 5) cloud resource. The
164 curated content was advertised on the main hackathon website. One dedicated channel
165 was created on the hackathon Discord server. Individuals could determine the nature of
166 their experiences and the skills they liked to acquire. Participants could form their own

₁₆₇ study group and on any selected topic. We would like to continue the experimentation
₁₆₈ on this format in the coming year.

₁₆₉ 3 Platforms, website, and IT

₁₇₀ *Anibal Solon Heinsfeld*

₁₇₁

₁₇₂ Trying to bring a positive experience for both virtual and in-person attendees, we
₁₇₃ implemented several integrated solutions to ease communication in the different phases
₁₇₄ of the Hackathon, focusing on a single platform for the main event.

₁₇₅ The first solution was the project's advertisement, in which the community promotes
₁₇₆ their projects, the goals for the Hackathon, and relevant information to get people inter-
₁₇₇ ested and set to collaborate. To do so, we used the Github Issues feature in the Hackathon
₁₇₈ repository as the entrance for projects. Github Issues has been proven to be accepted by
₁₇₉ the community that relies on Github for code versioning, and was a successful approach
₁₈₀ in past hackathons.

₁₈₁ In this edition, we were able to use Github Issue forms, a beta feature in Github. Past
₁₈₂ use of issues for project registration relies on Markdown code to specify which information
₁₈₃ the hacker needs to provide. However, the code can be easily broken and changed, which
₁₈₄ makes it harder to parse the information in automated setups. Towards this issue, the
₁₈₅ Issue Form can lower the barrier when submitting a project. By specifying form fields
₁₈₆ for the participants to fill, they faced a common web form instead of a Markdown editor,
₁₈₇ bringing more structure to their inputs and not requiring them to write code. After
₁₈₈ the organisers' quick validation, the project information was provided to the rest of the
₁₈₉ system. Per an automated pipeline, this information was compiled into the website.

₁₉₀ The second solution was the central platform for real-time communication, namely
₁₉₁ Discord. For the first time using the platform for an OHBM Hackathon, Discord showed
₁₉₂ potential in bringing an all-in-one solution. Its track record with different communities
₁₉₃ and their formats was an essential prospect for the success of a hybrid hackathon, together

194 with the different ways of communicating provided by the platform. Specifically, Discord
195 offered chat and audio/video channels, with fine-tuned controls on permissions to see a
196 channel, speak and use the camera, and send messages. With these features, we were able
197 to create experiences for the attendants, such as text channels for consolidating inform-
198 ation about the hackathon, main stages controlled by the hub hosts, a channel to join
199 projects and hubs, and integrated text & voice channels for each project. The main stage
200 was connected to a laptop in the venue, providing synchronous streaming for announce-
201 ments, project pitches and progress reports for those participating virtually. The project
202 channels were automatically created together with the Github Issues. However, given
203 the thriving number of projects, the Discord server was replete with project channels.
204 Such a scenario was overwhelming for the attendants, especially for those approaching
205 Discord for the first time. To ameliorate this issue, a main projects channel was created,
206 so attendants could automatically join projects via related emoji reactions. The project
207 channels were of public access; however, only displayed upon joining the project. Besides
208 initial technical hiccups, the platform proved a good alternative for such an event format.

209 These integrated solutions smoothed the organisation of the event, the virtual platform
210 provided great support for the on-line participants. However, there was not a lot of inter-
211 action between in-person and online participants, and projects were mainly either virtual
212 or in-person (with few exceptions). This is probably because hybrid hacking provides chal-
213 lenges for organisation and attendants alike, even just in the physical limitations of being
214 able to have a video conference with a split team. It is important to consider, however,
215 that this was also the first in-person event for many participants, who preferred in-person
216 interaction and collaboration rather than the same on-line interaction that characterised
217 such events in the previous two years.

218 4 Project Reports

219 The peculiar nature of a Brainhack¹ reflects in the nature of the projects developed
220 during the event, that can span very different types of tasks. While most projects feature

more ‘hackathon-style’ software development, in the form of improving software integration (Section 4.4), API refactoring (Section 4.11), or creation of new toolboxes and platforms (Sections 4.9, 4.10 and 4.13), the inclusion of newcomers and participants with less strong software development skills can foster projects oriented to user testing (Sections 4.3 and 4.9) or documentation compilation (Section 4.12). The scientific scopes of Brainhacks were reflected in projects revolving around data exploration (Sections 4.1 and 4.7) or model development (Section 4.13), or adding aspects of open science practices (namely, the Brain Imaging Data Structure) to toolboxes (Sections 4.6 and 4.14). Finally, fostering a collaborative environment and avoiding pitching projects against each others not only opens up the possibility for participants to fluidly move between different groups, but also to have projects which sole aim is supporting other projects (Section 4.2), learning new skills with entertaining tasks (Section 4.5), or fostering discussions and conversations among participants to improve the adoption of open science practices (??).

Following are the 14 submitted reports of the 23 projects presented at project wrap-up during the OHBM Brainhack.

4.1 Exploring the AHEAD brains together

Alessandra Pizzuti, Sebastian Dresbach, Satrajit Ghosh, Katja Heuer, Roberto Toro, Pierre-Louis Bazin

238

4.1.1 Introduction

One of the long-standing goals of neuroanatomy is to compare the cyto- and myeloarchitecture of the human brain. The recently made available 3D whole-brain post-mortem data set provided by Alkemade and colleagues² includes multiple microscopy contrasts and 7-T quantitative multi-parameter MRI reconstructed at 200µm from two human brains. Through the co-registration across MRI and microscopy modalities, this data set provides a unique direct comparison between histological markers and quantitative MRI parameters for the same human brain. In this BrainHack project, we explored this dataset, focusing on: (i) data visualization in online open science platforms, (ii) data in-

²⁴⁸ tegration of quantitative MRI with microscopy, (iii) data analysis of cortical profiles from
²⁴⁹ a selected region of interest.

²⁵⁰ **4.1.2 Results**

²⁵¹ Visualization and annotation of large neuroimaging data sets can be challenging, in
²⁵² particular for collaborative data exploration. Here we tested two different infrastruc-
²⁵³ tures: BrainBox <https://brainbox.pasteur.fr/>, a web-based visualization and an-
²⁵⁴ notation tool for collaborative manual delineation of brain MRI data, see e.g. Heuer and
²⁵⁵ colleagues³, and Dandi Archive <https://dandiarchive.org/>, an online repository of mi-
²⁵⁶ croscopy data with links to Neuroglancer <https://github.com/google/neuroglancer>.
²⁵⁷ While Brainbox could not handle the high resolution data well, Neuroglancer visualization
²⁵⁸ was successful after conversion to the Zarr microscopy format (Figure 1A).

²⁵⁹ To help users explore the original high-resolution microscopy sections, we also built a
²⁶⁰ python notebook to automatically query the stains around a given MNI coordinate using
²⁶¹ the Nighres toolbox⁴ (Figure 1B).

²⁶² For the cortical profile analysis we restricted our analysis on S1 (BA3b) as a part of
²⁶³ the somato-motor area from one hemisphere of an individual human brain. S1 is rather
²⁶⁴ thin (\sim 2mm) and it has a highly myelinated layer 4 (see arrow Figure 1C). In a future
²⁶⁵ step, we are aiming to characterize differences between S1 (BA3b) and M1 (BA4). For
²⁶⁶ now, we used the MRI-quantitative-R1 contrast to define, segment the region of interest
²⁶⁷ and compute cortical depth measurement. In ITK-SNAP⁵ we defined the somato-motor
²⁶⁸ area by creating a spherical mask (radius 16.35mm) around the ‘hand knob’ in M1. To
²⁶⁹ improve the intensity homogeneity of the qMRI-R1 images, we ran a bias field correction
²⁷⁰ (N4BiasFieldCorrection,⁶). Tissue segmentation was restricted to S1 and was obtained
²⁷¹ by combining four approaches: (i) fsl-fast⁷ for initial tissues probability map, (ii) semi-
²⁷² automatic histogram fitting in ITK-SNAP, (iii) Segmentator⁸, and (iv) manual editing.
²⁷³ We used the LN2_LAYERS program from LAYNII open source software⁹ to compute
²⁷⁴ the equi-volume cortical depth measurements for the gray matter. Finally, we evaluated
²⁷⁵ cortical depth profiles for three quantitative MRI contrasts (R1, R2, proton density)

276 and three microscopy contrasts (thionin, bieloschowsky, parvalbumin) by computing a
277 voxel-wise 2D histogram of image intensity (Figure 1C). Some challenges are indicated
278 by arrows 2 and 3 in the lower part of Figure 1C.

279 From this Brainhack project, we conclude that the richness of the data set must be
280 exploited from multiple points of view, from enhancing the integration of MRI with mi-
281 croscopy data in visualization software to providing optimized multi-contrast and multi-
282 modality data analysis pipeline for high-resolution brain regions.

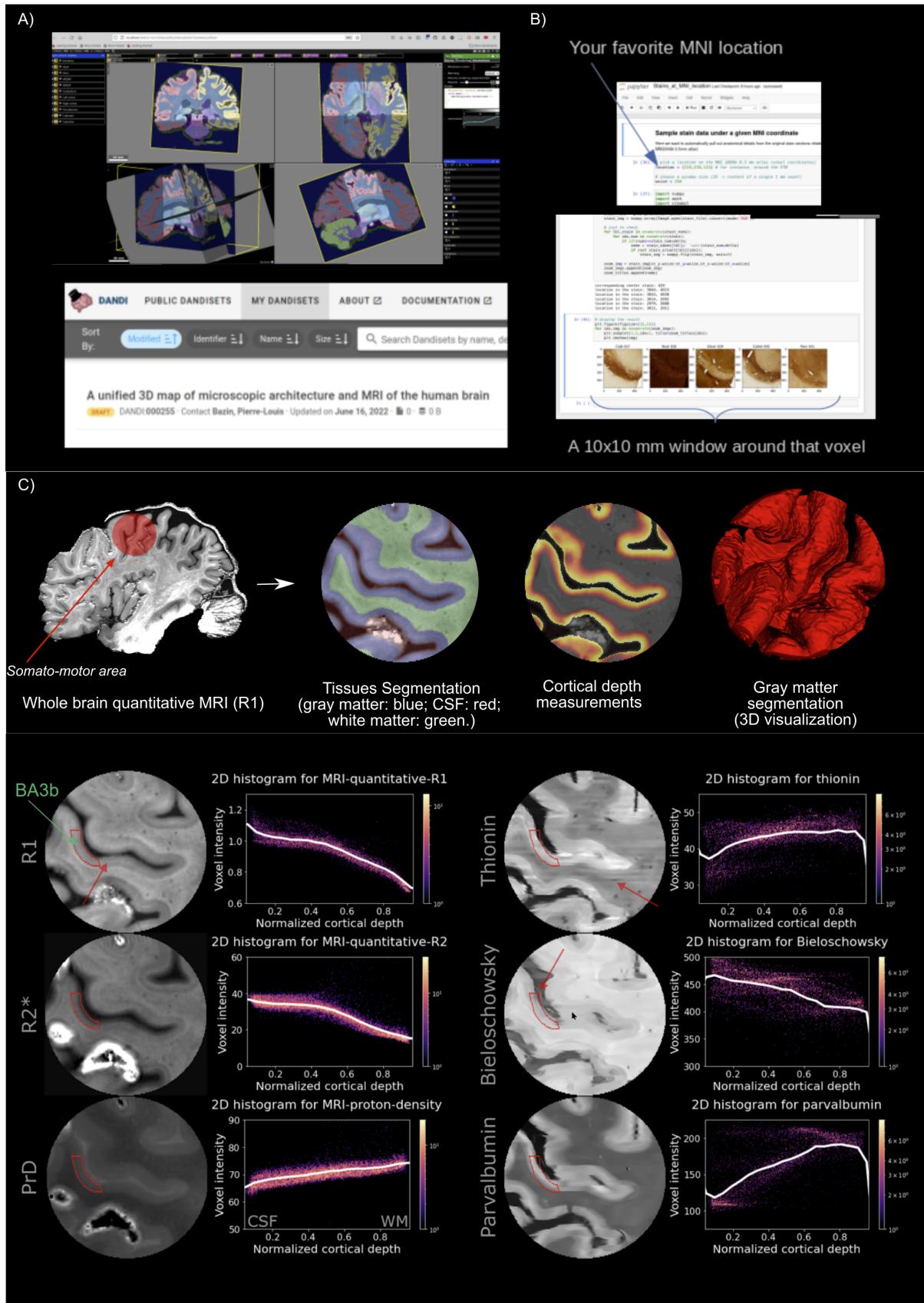


Figure 1: A) Neuroglancer visualization, B) section query notebook, C) Cortical ROI and corresponding depth histograms extracted from the different contrasts available.

283 **4.2 Brainhack Cloud**

284 *Steffen Bollmann, Isil Poyraz Bilgin, Peer Herholz, Rémi Gau, Samuel Guay, Johanna Bayer*

285

286 Today's neuroscientific research deals with vast amounts of electrophysiological, neuroima-
287 ging and behavioural data. The progress in the field is enabled by the widespread avail-
288 ability of powerful computing and storage resources. Cloud computing in particular offers
289 the opportunity to flexibly scale resources and it enables global collaboration across insti-
290 tutions. However, cloud computing is currently not widely used in the neuroscience field,
291 although it could provide important scientific, economical, and environmental gains con-
292 sidering its effect in collaboration and sustainability^{10,11}. One problem is the availability
293 of cloud resources for researchers, because Universities commonly only provide on-premise
294 high performance computing resources. The second problem is that many researchers lack
295 the knowledge on how to efficiently use cloud resources. This project aims to address both
296 problems by providing free access to cloud resources for the brain imaging community
297 and by providing targeted training and support.

298 A team of brainhack volunteers ([https://brainhack.org/brainhack_cloud/admins/
299 team/](https://brainhack.org/brainhack_cloud/admins/team/)) applied for Oracle Cloud Credits to support open-source projects in and around
300 brainhack with cloud resources. The project was generously funded by Oracle Cloud for
301 Research¹² with \$230,000.00 AUD from the 29th of January 2022 until the 28th of Janu-
302 ary 2024. To facilitate the uptake of cloud computing in the field, the team built several
303 resources (https://brainhack.org/brainhack_cloud/tutorials/) to lower the entry
304 barriers for members of the Brainhack community.

305 During the OHBM 2022 Brainhack, the team gave a presentation to share the cap-
306 abilities that cloud computing offers to the Brainhack community, how they can place
307 their resource requests and where they can get help. In total 11 projects were onboarded
308 to the cloud and supported in their specific use cases: One team utilised the latest GPU
309 architecture to take part in the Anatomical Tracings of Lesions After Stroke Grand Chal-
310 lenge. Others developed continuous integration tests for their tools using for example a
311 full Slurm HPC cluster in the cloud to test how their tool behaves in such an environ-

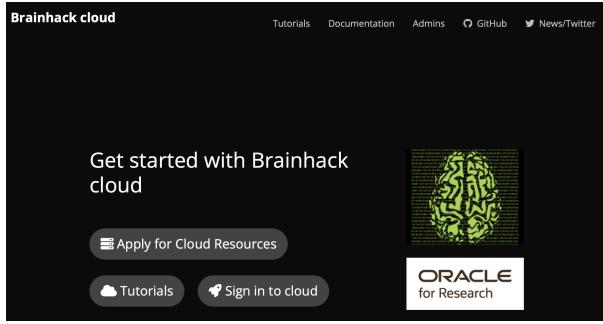


Figure 2: A team of brainhack volunteers, applied for Oracle Cloud Credits to support open source projects in and around brainhack with powerful cloud resources on the Oracle Cloud: https://brainhack.org/brainhack_cloud/

³¹² ment. Another group deployed the Neurodesk.org¹³ project on a Kubernetes cluster to
³¹³ make it available for a student cohort to learn about neuroimage processing and to get
³¹⁴ access to all neuroimaging tools via the browser. All projects will have access to these
³¹⁵ cloud resources until 2024 and we are continuously onboarding new projects onto the
³¹⁶ cloud (https://brainhack.org/brainhack_cloud/docs/request/).

³¹⁷ The Brainhack Cloud team plans to run a series of training modules in various Brain-
³¹⁸ hack events throughout the year to reach researchers from various backgrounds and in-
³¹⁹ crease their familiarity with the resources provided for the community while providing
³²⁰ free and fair access to the computational resources. The training modules will cover how
³²¹ to use and access computing and storage resources (e.g., generating SSH keys), to more
³²² advanced levels covering the use of cloud native technology like software containers (e.g.,
³²³ Docker/Singularity), container orchestration (e.g., Kubernetes), object storage (e.g, S3),
³²⁴ and infrastructure as code (e.g., Terraform).

³²⁵ 4.3 DataLad Catalog

³²⁶ Stephan Heunis, Adina S. Wagner, Alexander Q. Waite, Benjamin Poldrack, Christian Mönch, Julian Kosciessa,
³²⁷ Laura Waite, Leonardo Muller-Rodriguez, Michael Hanke, Michał Szczepanik, Remi Gau, Yaroslav O. Halchenko

³²⁸
³²⁹ The importance and benefits of making research data Findable, Accessible, Interoper-
³³⁰ able, and Reusable are clear¹⁴. But of equal importance is our ethical and legal obligations
³³¹ to protect the personal data privacy of research participants. So we are struck with this

332 apparent contradiction: how can we share our data openly... yet keep it secure and pro-
333 tected?

334 To address this challenge: structured, linked, and machine-readable metadata presents
335 a powerful opportunity. Metadata provides not only high-level information about our re-
336 search data (such as study and data acquisition parameters) but also the descriptive as-
337 pects of each file in the dataset: such as file paths, sizes, and formats. With this metadata,
338 we can create an abstract representation of the full dataset that is separate from the ac-
339 tual data content. This means that the content can be stored securely, while we openly
340 share the metadata to make our work more FAIR.

341 In practice, the distributed data management system DataLad¹⁵ and its extensions
342 for metadata handling and catalog generation are capable of delivering such solutions.
343 `datalad` (github.com/datalad/datalad) can be used for decentralised management of
344 data as lightweight, portable and extensible representations. `datalad-metlad` ([github.](https://github.com/datalad/datalad-metlad)
345 [com/datalad/datalad-metlad](https://github.com/datalad/datalad-metlad)) can extract structured high- and low-level metadata
346 and associate it with these datasets or with individual files. And at the end of the work-
347 flow, `datalad-catalog` (github.com/datalad/datalad-catalog) can turn the struc-
348 tured metadata into a user-friendly data browser.

349 This hackathon project focused on the first round of user testing of the alpha ver-
350 sion of `datalad-catalog`, by creating the first ever user-generated catalog (https://jkosciessa.github.io/datalad_cat_test). Further results included a string of new
351 issues focusing on improving user experience, detailed notes on how to generate a catalog
352 from scratch, and code additions to allow the loading of local web-assets so that any
353 generated catalog can also be viewed offline.

355 4.4 DataLad-Dataverse integration

356 *Benjamin Poldrack, Jianxiao Wu, Kelvin Sarink, Christopher J. Markiewicz , Alexander Q. Waite , Eliana Nicolaisen-*
357 *Sobesky, Shammi More, Johanna Bayer, Jan Ernsting, Adina S. Wagner, Roza G. Bayrak , Laura K. Waite, Michael*
358 *Hanke, Nadine Spychala*

359

360 The FAIR principles¹⁴ advocate to ensure and increase the Findability, Accessibil-
361 ity, Interoperability, and Reusability of research data in order to maximize their impact.
362 Many open source software tools and services facilitate this aim. Among them is the Data-
363 verse project¹⁶. Dataverse is open source software for storing and sharing research data,
364 providing technical means for public distribution and archival of digital research data, and
365 their annotation with structured metadata. It is employed by dozens of private or public
366 institutions worldwide for research data management and data publication. DataLad¹⁵,
367 similarly, is an open source tool for data management and data publication. It provides
368 Git- and git-annex based data versioning, provenance tracking, and decentral data dis-
369 tribution as its core features. One of its central development drivers is to provide stream-
370 lined interoperability with popular data hosting services to both simplify and robustify
371 data publication and data consumption in a decentralized research data management
372 system¹⁷. Past developments include integrations with the open science framework¹⁸ or
373 webdav-based services such as sciebo, nextcloud, or the European Open Science Cloud¹⁹.

374 In this hackathon project, we created a proof-of-principle integration of DataLad
375 with Dataverse in the form of the Python package `datalad-dataverse` ([github.com/
376 datalad/datalad-dataverse](https://github.com/datalad/datalad-dataverse)). From a technical perspective, main achievements include
377 the implementation of a git-annex special remote protocol for communicating with Data-
378 verse instances, a new `create-sibling-dataverse` command that is added to the DataLad
379 command-line and Python API by the `datalad-dataverse` extension, and standard re-
380 search software engineering aspects of scientific software such as unit tests, continuous
381 integration, and documentation.

382 From a research data management and user perspective, this development equips
383 DataLad users with the ability to programatically create Dataverse datasets (containers
384 for research data and their metadata on Dataverse) from DataLad datasets (DataLad's
385 Git-repository-based core data structure) in different usage modes. Subsequently, DataLad
386 dataset contents, its version history, or both can be published to the Dataverse dataset via
387 a 'datalad push' command. Furthermore, published DataLad datasets can be consumed
388 from Dataverse with a `datalad clone` call. A mode parameter configures whether Git

389 version history, version controlled file content, or both are published and determines
390 which of several representations the Dataverse dataset takes. A proof-of-principle imple-
391 mentation for metadata annotation allows users to supply metadata in JSON format, but
392 does not obstruct later or additional manual metadata annotation via Dataverse's web
393 interface.

394 Overall, this project delivered the groundwork for further extending and streamlining
395 data deposition and consumption in the DataLad ecosystem. With DataLad-Dataverse
396 interoperability, users gain easy additional means for data publication, archival, distribu-
397 tion, and retrieval. Post-Brainhack development aims to mature the current alpha version
398 of the software into an initial v0.1 release and distribute it via standard Python package
399 indices.

400 4.5 Exploding brains in Julia

401 *Ömer Faruk Gülbán, Leonardo Müller-Rodriguez*

402

403 Particle simulations are used to generate visual effects (in movies, games, etc.). In this
404 project, we explore how we can use magnetic resonance imaging (MRI) data to generate
405 interesting visual effects by using (2D) particle simulations. [Aside from providing an](#)
406 [entertaining avenue to the interested participants, our project has further educational](#)
407 [utility.](#) For instance, anatomical MRI data analysis is done in two major frameworks:
408 (1) manipulating fixed regularly spaced points in space (also known as Eulerian point of
409 view), and (2) manipulating moving irregularly spaced points in space (Lagrangian point
410 of view). For instance, bias field correction is commonly done from Eulerian point of view
411 (e.g. computing a bias field is similar to computing a particle velocity field in each frame
412 of the explosions), whereas cortical surface inflation is commonly done from Lagrangian
413 point of view of the MRI data (e.g. computing the inflated brain surface is similar to
414 computing the new positions of particles in each frame of the explosion). Therefore, our
415 project provides an educational opportunity for those who would like to peek into the
416 deep computational and data structure manipulation aspects of MRI image analysis.

417 We note that we already made two hackathon projects in 2020 (see below) and were first
418 inspired by a blog post (https://nialltl.neocities.org/articles/mpm_guide.html)
419 on the material point method^{20,21,22}. Our additional aim in Brainhack 2022 is to convert
420 our previous progress in Python programming language to Julia. The reason why we
421 have moved to Julia language is because we wanted to explore this new programming
422 language's potential for developing MRI image analysis methods as it has convenient
423 parallelization methods that speeds-up the particle simulations (and any other advanced
424 image manipulation algorithms).

425

426 Our previous efforts are documented at:

- 427 1. 2020 OpenMR Benelux: [https://github.com/OpenMRBenelux/openmrb2020-hackathon/](https://github.com/OpenMRBenelux/openmrb2020-hackathon/issues/7)
428 issues/7
- 429 2. 2020 OHBM Brainhack: <https://github.com/ohbm/hackathon2020/issues/124>
- 430 3. Available within the following github repository: <https://github.com/ofgulban/slowest-particle-simulator-on-earth>
-

432

433 As a result of this hackathon project, we delivered a video compilation of our anim-
434 ations (Figure 3) which can be seen at https://youtu.be/_5ZDctWv5X4. We highlight
435 that in addition to its educational value, our project provided stress relief by means of
436 entertaining the participants after the pandemic. We believe that our project provides
437 a blueprint for the future brainhacks where MRI science, computation, and education
438 can be disseminated within an engaging and entertaining context. Our future efforts will
439 involve sophisticating the particle simulations, the initial simulation parameters to gen-
440 erate further variations of the visual effects, and potentially synchronizing the simulation
441 effects with musical beats.

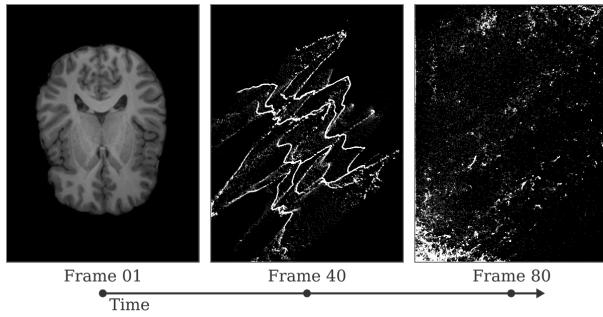


Figure 3: A video compilation of brain explosions can be seen at https://youtu.be/_5ZDctWv5X4.

442 4.6 FLUX: A pipeline for MEG analysis and beyond

443 *Oscar Ferrante, Tara Ghafari, Ole Jensen*

444
445 FLUX²³ is an open-source pipeline for analysing magnetoencephalography (MEG)
446 data. There are several toolboxes developed by the community to analyse MEG data.
447 While these toolboxes provide a wealth of options for analyses, the many degrees of free-
448 dom pose a challenge for reproducible research. The aim of FLUX is to make the analyses
449 steps and setting explicit. For instance, FLUX includes the state-of-the-art suggestions
450 for noise cancellation as well as source modelling including pre-whitening and handling
451 of rank-deficient data.

452 So far, the FLUX pipeline has been developed for MNE-Python²⁴ and FieldTrip²⁵
453 with a focus on the MEGIN/Elekta system and it includes the associated documents as
454 well as codes. The long-term plan for this pipeline is to make it more flexible and versatile
455 to use. One key motivation for this is to facilitate open science with the larger aim of
456 fostering the replicability of MEG research.

457 These goals can be achieved in mid-term objectives, such as making the FLUX pipeline
458 fully BIDS compatible and more automated. Another mid-term goal is to containerize
459 the FLUX pipeline and the associated dependencies making it easier to use. Moreover,
460 expanding the applications of this pipeline to other systems like MEG CTF, Optically
461 Pumped Magnetometer (OPM) and EEG will be another crucial step in making FLUX
462 a more generalized neurophysiological data analysis pipeline.

463 During the 2022 Brainhack, the team focused on incorporating the BIDS standard into

⁴⁶⁴ the analysis pipeline using MNE_BIDS²⁶. Consequently, an updated version of FLUX
⁴⁶⁵ was released after the Brainhack meeting.

⁴⁶⁶ **4.7 Evaluating discrepancies in hippocampal segmentation pro-**
⁴⁶⁷ **tocols using automatic prediction of MRI quality (MRIQC)**

⁴⁶⁸ *Jacob Sanz-Robinson, Mohammad Torabi, Tyler James Wishard*

⁴⁶⁹

⁴⁷⁰ **4.7.1 Introduction**

⁴⁷¹ Neuroimaging study results can vary significantly depending on the processing pipelines
⁴⁷² utilized by researchers to run their analyses, contributing to reproducibility issues. Re-
⁴⁷³ searchers in the field are often faced with multiple choices of pipelines featuring similar
⁴⁷⁴ capabilities, which may yield different results when applied to the same data^{27,28}. While
⁴⁷⁵ these reproducibility issues are increasingly well-documented in the literature, there is
⁴⁷⁶ little existing research explaining why this inter-pipeline variability occurs or the factors
⁴⁷⁷ contributing to it. In this project, we set out to understand what data-related factors
⁴⁷⁸ impact the discrepancy between popular neuroimaging processing pipelines.

⁴⁷⁹ **4.7.2 Method**

⁴⁸⁰ The hippocampus is a structure commonly associated with memory function and
⁴⁸¹ dementia, and the left hippocampus is proposed to have higher discriminative power for
⁴⁸² identifying the progression of Alzheimer's disease than the right hippocampus in multiple
⁴⁸³ studies²⁹. We obtained left hippocampal volumes using three widely-used neuroimaging
⁴⁸⁴ pipelines: FSL 5.0.9³⁰, FreeSurfer 6.0.0³¹, and ASHS 2.0.0 PMC-T1 atlas³². We ran the
⁴⁸⁵ three pipelines on T1 images from 15 subjects from the Prevent-AD Alzheimer's dataset³³,
⁴⁸⁶ composed of cognitively healthy participants between the ages of 55-88 years old that
⁴⁸⁷ are at risk of developing Alzheimer's Disease. We ran MRIQC³⁴ - a tool for performing
⁴⁸⁸ automatic quality control and extracting quality measures from MRI scans - on the 15
⁴⁸⁹ T1 scans and obtained Image Quality Metrics (IQMs) from them. We then found the

⁴⁹⁰ correlations between the IQMs and the pairwise inter-pipeline discrepancy of the left
⁴⁹¹ hippocampal volumes for each T1 scan.

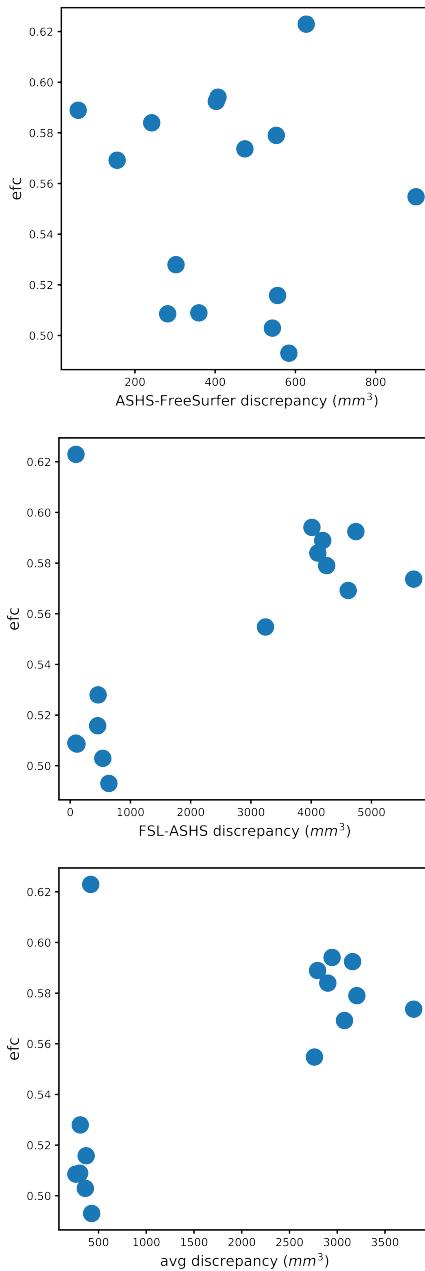


Figure 4: Plots showing the association between left hippocampal volume discrepancies and MRIQC's EFC quality measure for each of the pipeline pairings.

⁴⁹² **4.7.3 Results**

⁴⁹³ We found that for The FSL-FreeSurfer and FSL-ASHs discrepancies, MRIQC's EFC
⁴⁹⁴ measure produced the highest correlation, of 0.69 and 0.64, respectively. The EFC “uses

495 the Shannon entropy of voxel intensities as an indication of ghosting and blurring induced
496 by head motion”³⁵. No such correlations were found for the ASHS-FreeSurfer discrepan-
497 cies. Figure 4 shows a scatter plot of the discrepancies in left hippocampal volume and
498 EFC IQM for each pipeline pairing. The preliminary results suggest that FSL’s hippo-
499 campal segmentation may be sensitive to head motion in T1 scans, leading to larger
500 result discrepancies, but we require larger sample sizes to make meaningful conclusions.
501 The code for our project can be found on GitHub at this link.

502 **4.7.4 Conclusion and Next Steps**

503 In this project, we investigated the correlation between MRIQC’s IQMs and discrep-
504 ances in left hippocampal volume derived from three common neuroimaging pipelines on
505 15 subjects from the Prevent-AD study dataset. While our preliminary results indicate
506 image ghosting and blurring induced by head motion may play a role in inter-pipeline
507 result discrepancies, the next steps of the project will consist of computing the corre-
508 lations on the full 308 subjects of the Prevent-AD dataset to investigate whether they
509 persist with the full sample.

510 **4.8 Accelerating adoption of metadata standards for dataset descriptors**

511 *Cassandra Gould van Praag, Felix Hoffstaedter, Sebastian Urchs*

512

513 We have used the space of the brainhack to discuss challenges that are hindering
514 wide adoption of metadata standards in the neuroimaging community and to brainstorm
515 possible solutions to accelerate it. Although our project was conceptual and we did not
516 develop any tools during the project, the outcome of our discussions have directly influ-
517 enced the development of tools such as neurobagel³⁶ (<https://neurobagel.org/>) after
518 the brainhack.

519 Thanks to efforts of the neuroimaging community, not least the brainhack community¹,
520 datasets are increasingly shared on open data repositories like OpenNeuro³⁷ using stand-
521 ards like BIDS³⁸ for interoperability. As the amount of datasets and data repositories

522 increases, we need to find better ways to search across them for samples that fit our
523 research questions. In the same way that the wide adoption of BIDS makes data sharing
524 and tool development easier, the wide adoption of consistent vocabulary for demographic,
525 clinical and other sample metadata would make data search and integration easier. We
526 imagine a future platform that allows cross dataset search and the pooling of data across
527 studies. Efforts to establish such metadata standards have had some success in other
528 communities^{39,40}, but adoption in the neuroscience community so far has been slow.

529 We believe that an important social challenge for the wider adoption of metadata
530 standards is that it is hard to demonstrate their value without a practical use case. We
531 therefore think that rather than focusing on building better standards, in the short term
532 we need to prioritize small, but functional demonstrations that help convey the value of
533 these standards and focus on usability and ease of adoption. Having consistent names
534 and format for even a few metadata variables like age, sex, and diagnosis already allows
535 for interoperability and search across datasets. Selecting a single vocabulary that must
536 be used for annotating e.g. diagnosis necessarily lacks some precision but avoids the need
537 to align slightly different versions of the same terms. Accessible tools can be built to
538 facilitate the annotation process of such a basic metadata standard. The best standard
539 will be poorly adopted if there are no easy to use tools that implement it. Efforts like the
540 neurobagel project (<https://neurobagel.org/>) are trying to implement this approach
541 to demonstrate a simple working use case for cross dataset integration and search. Our
542 goal is to use such simpler demonstrations to build awareness and create a community
543 around the goal of consistent metadata adoption.

544 Our long term goal is to use the awareness of the value of shared metadata standards
545 to build a community to curate the vocabularies used for annotation. The initially small
546 number of metadata variables will have to be iteratively extended through a community
547 driven process to determine what fields should be standardized to serve concrete use cases.
548 Rather than creating new vocabularies the goal should be to curate a list of existing ones
549 that can be contributed to where terms are inaccurate or missing. The overall goal of
550 such a community should be to build consensus on and maintain shared standards for

551 the annotation of neuroimaging metadata that support search and integration of data for
552 an ever more reproducible and generalizable neuroscience.

553 **4.9 The NARPS Open Pipelines Project**

554 *Elodie Germani, Arshitha Basavaraj, Trang Cao, Rémi Gau, Anna Menacher, Camille Maumet*

555

556 The goal of the NARPS Open Pipelines Project is to provide a public codebase that
557 reproduces the 70 pipelines chosen by the 70 teams of the NARPS study⁴¹. The project is
558 public and the code hosted on GitHub at [https://github.com/Inria-Empenn/narps_](https://github.com/Inria-Empenn/narps_open_pipelines)
559 [open_pipelines](#).

560 This project initially emerged from the idea of creating an open repository of fMRI
561 data analysis pipelines (as used by researchers in the field) with the broader goal to
562 study and better understand the impact of analytical variability. NARPS – a many-
563 analyst study in which 70 research teams were asked to analyze the same fMRI dataset
564 with their favorite pipeline – was identified as an ideal usecase as it provides a large array
565 of pipelines created by different labs. In addition, all teams in NARPS provided extensive
566 (textual) description of their pipelines using the COBIDAS⁴² guidelines. All resulting
567 statistic maps were shared on NeuroVault⁴³ and can be used to assess the success of the
568 reproductions.

569 At the OHBM Brainhack 2022, our goal was to improve the accessibility and reusab-
570 ility of the database, to facilitate new contributions and to reproduce more pipelines. We
571 focused our efforts on the first two goals. By trying to install the computing environ-
572 ment of the database, contributors provided feedback on the instructions and on specific
573 issues they faced during the installation. Two major improvements were made for the
574 download of the necessary data: the original fMRI dataset and the original results (stat-
575 istic maps stored in NeuroVault) were added as submodules to the GitHub repository.
576 Finally, propositions were made to facilitate contributions: the possibility to use of the
577 Giraffe toolbox⁴⁴ for contributors that are not familiar with NiPype⁴⁵ and the creation
578 of a standard template to reproduce a new pipeline.

579 With these improvements, we hope that it will be easier for new people to contribute
580 to reproduction of new pipelines. We hope to continue growing the codebase in the future.

581 **4.10 NeuroCausal: Development of an Open Source Platform**
582 **for the Storage, Sharing, Synthesis, and Meta-Analysis of**
583 **Neuropsychological Data**

584 *Isil Poyraz Bilgin, Francois Paugam, Ruoqi Huang, Ana Luísa Pinho, Yuchen Zhou, Sladjana Lukic, Pedro Pinheiro-*

585 *Chagas, Valentina Borghesani*

586

587 Cognitive neuroscience has witnessed great progress since modern neuroimaging em-
588 braced an open science framework, with the adoption of shared principles¹⁴, standards³⁸,
589 and ontologies⁴⁶, as well as practices of meta-analysis^{47,48} and data sharing⁴³. However,
590 while functional neuroimaging data provide correlational maps between cognitive func-
591 tions and activated brain regions, its usefulness in determining causal link between specific
592 brain regions and given behaviors or functions is disputed^{49,50}. On the contrary, neuropsy-
593 chological data enable causal inference, highlighting critical neural substrates and opening
594 a unique window into the inner workings of the brain⁵¹. Unfortunately, the adoption of
595 Open Science practices in clinical settings is hampered by several ethical, technical, eco-
596 nomic, and political barriers, and as a result, open platforms enabling access to and
597 sharing clinical (meta)data are scarce⁵².

598 With our project, NeuroCausal (<https://neurocausal.github.io/>), we aim to build
599 an online platform and community that allows open sharing, storage, and synthesis of
600 clinical (meta) data crucial for the development of modern, transdiagnostic, accessible,
601 and replicable (i.e., FAIR: Findability, Accessibility, Interoperability, and Reusability)
602 neuropsychology. The project is organized into two infrastructural stages: first, published
603 peer-reviewed papers will be scrapped to collect already available (meta)data; second,
604 our platform will allow direct uploading of clinical (de-identified) brain maps and their
605 corresponding metadata.

606 The meta-analysis pipeline developed for the first stage of the project is inspired by
607 and built upon the functionalities of NeuroQuery⁴⁷, a successful large-scale neuroimaging
608 meta-analytic platform. The first stage is the development of the code base allowing (1)

609 downloading and filtering of neuropsychological papers, (2) extraction of reported brain
610 lesion locations and their conversion into a common reference space (3) extraction of
611 clinical and behavioral symptoms and their translation into a common annotation scheme,
612 (4) learning the causal mapping between the neural and neuropsychological information
613 gathered.

614 The second stage of the study aims at creating an online platform that allows for the
615 direct uploading of clinical brain maps and their corresponding metadata. The platform
616 will provide a basic automated preprocessing and a data-quality check pipeline, ensuring
617 that all the ethical norms regarding patient privacy are met. The platform will automatic-
618 ally extract and synthesize key data to ultimately create probabilistic maps synthesizing
619 transdiagnostic information on symptom-structure mapping, which will be dynamically
620 updated as more data are gathered.

621 The nature of the project requires expertise in different fields (from clinical neuros-
622 cience to computer science) in order to overcome both technical and theoretical chal-
623 lenges. The OHBM Brainhack 2022 gave us the opportunity to set the first stones. In
624 small subteams, we worked on developing three key building blocks: (1) the input filter-
625 ing pipeline to ensure the downloaded papers are neuropsychological in nature and offer
626 causal symptom-structure mapping; (2) the extraction of key terms occurrences in the
627 text as to assess which neural space is reported (as they will have to be converted to a
628 common one), (3) the curation of clinical ontology mapping specific neuropsychological
629 batteries and tasks to the cognitive term(s) they touch upon. [During the hackathon we](#)
630 [worked on developing three key building blocks in small subteams. First, we prepared an](#)
631 [input filtering pipeline to ensure that the downloaded papers are neuropsychological in](#)
632 [nature \(and thus offer causal symptom-structure mapping\): we count the occurrences of](#)
633 [clinically relevant terms, and papers are included only if they pass an arbitrary threshold.](#)
634 [Second, we coded a script automatically returning for each paper information on the](#)
635 [neural spaced used \(e.g., which atlas? MNI coordinates?\), a crucial step to enable future](#)
636 [conversion to a common reference space. Finally, we curated a list of clinically relevant](#)
637 [terms and constructs \(a clinical ontology\) that maps specific neuropsychological batteries](#)

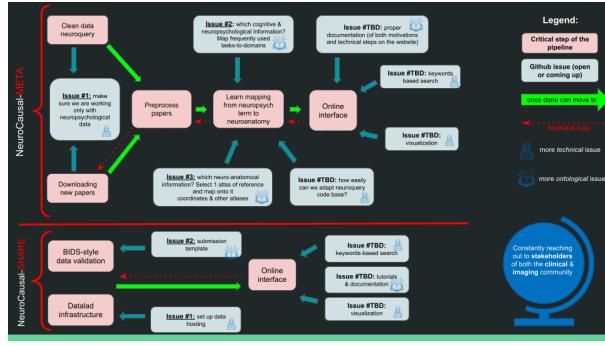


Figure 5: NeuroCausal: The future of neuropsychology, i.e. brain lesions-symptom mapping, will be transdiagnostic, open, and FAIR: we set out to provide the field with an open-source platform fostering storage, sharing, synthesis, and meta-analysis of clinical data.

638 and tasks to the cognitive term(s) they touch upon.

639 As we keep tackling our roadmap (Figure 1), we believe our efforts will help promote
 640 open science practices in clinical neuroscience to the benefit of both the neuroscientific
 641 and the clinical communities.

642 **Acknowledgments :** The authors would like to thank Eric Earl, Samuel Guay,
 643 Jerome Dockès, Bertrand Thirion, Jean Baptiste Poline, Yaroslav Halchenko, Sara El-
 644 Gebali and the whole Open Life Science team for their help and support.

645 4.11 Neuroscout: A platform for fast and flexible re-analysis of 646 (naturalistic) fMRI studies

647 *Alejandro De La Vega, Roberta Rocca, Sam Nastase, Peer Horholz, Jeff Menth, Kevin Sitek, Caroline Martin, Leonardo Muller-Rodriguez, Kan Keeratimahat, Dylan Nielson*

649

650 Neuroscout is an end-to-end platform for analysis of naturalistic fMRI data designed to
 651 facilitate the adoption of robust and generalizable research practices. Neuroscout's goal is
 652 to make it easy to analyze complex naturalistic fMRI datasets by providing an integrated
 653 platform for model specification and automated statistical modeling, reducing technical
 654 barriers. Importantly, Neuroscout is at its core a platform for reproducible analysis of
 655 fMRI data in general, and builds upon a set of open standards and specifications to
 656 ensure analyses are Findable, Accessible, Interoperable, and Reusable (FAIR).

657 In the OHBM Hackathon, we iterated on several important projects that substantially
658 improved the general usability of the Neuroscout platform. First, we launched a revamped
659 and unified documentation which links together all of the subcomponents of the Neur-
660 oscout platform (<https://neuroscout.github.io/neuroscout-docs/>). Second, we fa-
661 cilitated access to Neuroscout’s data sources by simplifying the design of Python API,
662 and providing high-level utility functions for easy programmatic data queries. Third, we
663 updated a list of candidate naturalistic and non-naturalistic datasets amenable for index-
664 ing by the Neuroscout platform, ensuring the platform stays up to date with the latest
665 public datasets.

666 In addition, important work was done to expand the types of analyses that can be
667 performed with naturalistic data in the Neuroscout platform. Notably, progress was made
668 in integrating Neuroscout with Himalaya, a library for efficient voxel wide encoding mod-
669 eling with support for banded penalized regression. In addition, a custom image-on-scalar
670 analysis was prototyped on naturalistic stimuli via the publicly available naturalistic fea-
671 tures available in the Neuroscout API. Finally, we also worked to improve documentation
672 and validation for BIDS StatsModels, a specification for neuroimaging statistical models
673 which underlies Neuroscout’s automated model fitting pipeline.

674 4.12 Physiopy - Documentation of Physiological Signal Best Prac- 675 tices

676 *Sarah E. Goodale, Ines Esteves, Roza G. Bayrak, Neville Magielse, Stefano Moia, Yu-Fang Yang, The Physiopy*
677 *Community*

678

679 Physiological data provides a representation of a subject’s internal state with respect
680 to peripheral measures (i.e., heart rate, respiratory rate, etc.). Recording physiological
681 measures is key to gain understanding of sources of signal variance in neuroimaging data
682 that arise from outside of the brain⁵³. This has been particularly useful for functional
683 magnetic resonance imaging (fMRI) research, improving fMRI time-series model accuracy,
684 while also improving real-time methods to monitor subjects during scanning^{54,55}.

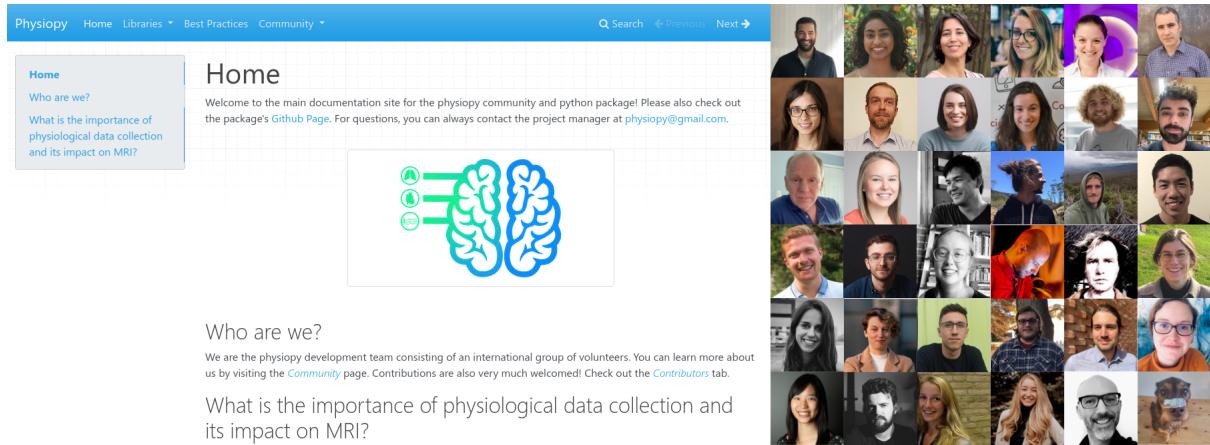


Figure 6: Left: Current version of the documentation homepage; Right: Physiopy Contributors

685 Physiopy (<https://github.com/physiopy>) is an open and collaborative community,
 686 formed around the promotion of physiological data collection and incorporation in neuroima-
 687 ging studies. Physiopy is focused on two main objectives. The first is the community-based
 688 development of tools for fMRI-based physiological processing. At the moment, there are
 689 three toolboxes: *phys2bids* (physiological data storage and conversion to BIDS format⁵⁶),
 690 *peakdet* (physiological data processing), and *phys2denoise* (fMRI denoising). The second
 691 objective is advancing the general knowledge of physiological data collection in fMRI
 692 by hosting open sessions to discuss best practices of physiological data acquisition, pre-
 693 processing, and analysis, and promoting community involvement. Physiopy maintains
 694 documentation with best practices guidelines stemming from these joint discussions and
 695 recent literature.

696 At the OHBM 2022 Brainhack, we aimed to improve our community documentation
 697 by expanding on best practices documentation, and gathering libraries of complementary
 698 open source software. This provides new users resources for learning about the process
 699 of physiological collection as well as links to already available resources. The short-term
 700 goal for the Brainhack was to prepare a common platform (and home) for our docu-
 701 mentation and repositories. We prioritised fundamental upkeep and content expansion,
 702 adopting Markdown documents and GitHub hosting to minimise barriers for new con-
 703 tributors. Over the course of the Brainhack, and with the joint effort within three hubs
 704 (Glasgow, EMEA and Americas), we were able to improve the current community web-

site by rethinking its structure and adding fundamental content relative to who we are, contributions, and updated best practices, such as creating home pages, easy to find and navigate contribution tabs, adding new information from community best practices discussions as well as links to relevant software and datasets. Additionally, we aggregated the information scattered across different repositories, allowing important information for both the community and new collaborators to be accessible in a single location.

The long-term goals of the community are to develop and sustain knowledge and instruments for physiological signal adoption in fMRI settings. Our aim is to facilitate the coming-together of researchers that are just starting to include physiological measures and experienced users. This community will then provide consensus guidelines for standardised data collection and preprocessing. Building on what we have already achieved, we will continue to promote and document best practices. Further development is ongoing and anyone that is interested in physiological signal collection for fMRI data, independently of their level and type of expertise, is highly encouraged to check Physiopy out, to join the community, or to contribute in any way.

4.13 Handling multiple testing problem through effect calibration: implementation using PyMC

Lea Waller, Kelly Garner, Christopher R. Nolan, Daniel Borek, Gang Chen

723

4.13.1 Introduction

Human brain imaging data is massively multidimensional, yet current approaches to modelling functional brain responses entail the application of univariate inferences to each voxel separately. This leads to the multiple testing problem and unrealistic assumptions about the data such as artificial dichotomization (statistically significant or not) in result reporting. The traditional approach of massively univariate analysis assumes that no information is shared across the brain, effectively making a strong prior assumption of a uniform distribution of effect sizes, which is unrealistic given the connectivity of the

⁷³² human brain. The consequent requirement for multiple testing adjustments results in the
⁷³³ *calibration of statistical evidence* without considering the estimation of effect, leading to
⁷³⁴ substantial information loss and an unnecessarily heavy penalty.

⁷³⁵ A more efficient approach to handling multiplicity focuses on the *calibration of effect*
⁷³⁶ *estimation* under a Bayesian multilevel modeling framework with a prior assumption of,
⁷³⁷ for example, normality across space⁵⁷. The methodology has previously been implemented
⁷³⁸ at the region level into the AFNI program RBA⁵⁸ using Stan through the R package brms⁵⁹.

⁷³⁹ We intend to achieve two goals in this project:

- ⁷⁴⁰ (i) To re-implement the methodology using PyMC improve the performance and flexi-
⁷⁴¹ bility of the modeling approach.
- ⁷⁴² (ii) To explore the possibility of analyzing voxel-level data using the multilevel modeling
⁷⁴³ approach

⁷⁴⁴ 4.13.2 Implementation using PyMC

⁷⁴⁵ We used the dataset from Chen and colleagues⁵⁷ to validate our PyMC implement-
⁷⁴⁶ ation. The data contain the subject-level response variable y and a predictor of the
⁷⁴⁷ behavioral measure x from $S = 124$ subjects at $R = 21$ regions. The modeling framework
⁷⁴⁸ is formulated for the data y_{rs} of the s th subject at the r th region as below,

$$\begin{aligned} y_{rs} &\sim \mathcal{N}(\mu_{rs}, \sigma^2) \\ \mu_{rs} &= \alpha_0 + \alpha_1 x_s + \theta_{0r} + \theta_{1r} x_s + \eta_s \\ \begin{bmatrix} \theta_{0r} \\ \theta_{1r} \end{bmatrix} &\sim \mathcal{N}(\mathbf{0}_{2 \times 1}, \mathbf{S}_{2 \times 2}) \\ \eta_s &\sim \mathcal{N}(0, \tau^2) \end{aligned} \tag{1}$$

where $r = 1, 2, \dots, R$ and $s = 1, 2, \dots, S$

⁷⁴⁹ In the model, μ_{rs} and σ are the mean effect and standard deviation of the s th subject
⁷⁵⁰ at the r th region, α_0 and α_1 are the overall mean and slope effect across all regions and

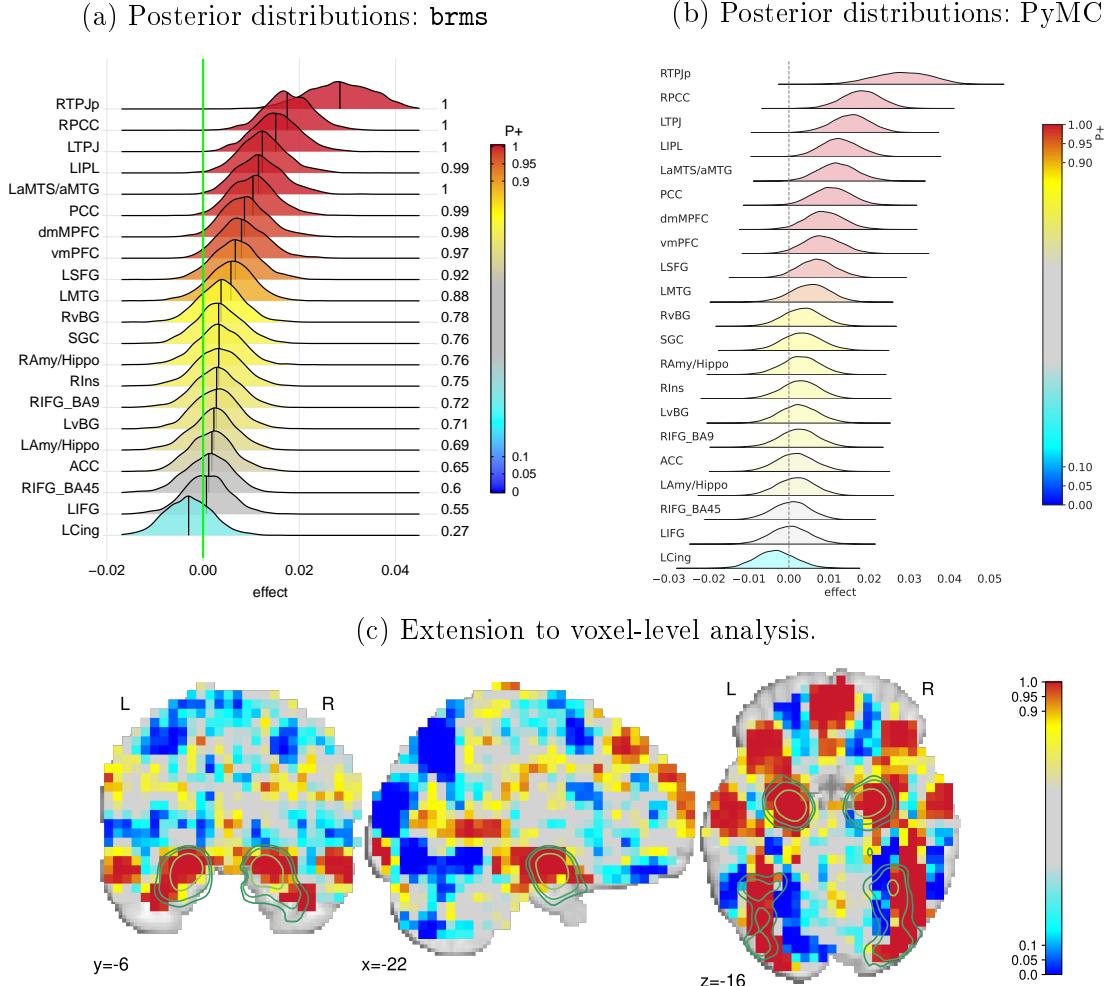


Figure 7: Validation of implementation using PyMC. (A) Posterior distributions of region-level behavior effects using `brms`. (B) Posterior distributions of region-level behavior effects using PyMC. (C) Posterior probabilities of the voxel-level effects being positive or negative, obtained using PyMC (plotted using Nilearn and overlaid in green with the NeuroQuery⁴⁷ map for the term “emotional faces”).

751 subjects, θ_{0r} and θ_{1r} are the mean and slope effect at the r th region, η_s is the mean effect
 752 of the s th subject, $\mathbf{S}_{2 \times 2}$ is the variance-covariance of the mean and slope effect at the r th
 753 region, and τ is the standard deviation of the s th subject’s effect η_s .

754 We implemented this model using the PyMC probabilistic programming framework⁶⁰,
 755 and the Bayesian Model-Building Interface (BAMBI)⁶¹. The latter is a high-level interface
 756 that allows for specification of multilevel models using the formula notation that is also
 757 adopted by `brms`. A notebook describing the implementation is available here. Our PyMC
 758 implementation was successfully validated: as shown in Figure 7a and Figure 7b, the
 759 posterior distributions from the PyMC implementation matched very well with their

760 counterparts from the `brms` output.

761 4.13.3 Extension of Bayesian multilevel modeling to voxel-level analysis

762 After exploring the model on the region level, we wanted to see if recent computational
763 and algorithmic advances allow us to employ the multilevel modeling framework on the
764 voxel level as well. We obtained the OpenNeuro dataset `ds000117`⁶² from an experiment
765 based on a face processing paradigm. Using `HALFpipe`⁶³, which is based on `fMRIPrep`⁶⁴,
766 the functional images were preprocessed with default settings and z -statistic images were
767 calculated for the contrast “famous faces + unfamiliar faces versus 2 · scrambled faces”.

768 We applied the same modeling framework and PyMC code as for region-based ana-
769 lysis, but without the explanatory variable x in the model (Equation (1)). To reduce
770 computational and memory complexity, the z -statistic images were downsampled to an
771 isotropic resolution of 5mm. Using the GPU-based `nuts_numpyro` sampler⁶⁵ with default
772 settings, we were able to draw 2,000 posterior samples of the mean effect parameter for
773 each of the 14,752 voxels. Sampling four chains took 23 minutes on four Nvidia Tesla
774 V100 GPUs.

775 The resulting posterior probabilities are shown in Figure 7c overlaid with the meta-
776 analytic map for the term “emotional faces” obtained from NeuroQuery⁴⁷. The posterior
777 probability map is consistent with meta-analytic results, showing strong statistical evid-
778 ence in visual cortex and amygdala voxels. The posterior probability maps also reveal
779 numerous other clusters of strong statistical evidence for both positive and negative ef-
780 ffects.

781 This implementation extension shows that large multilevel models are approaching
782 feasibility, suggesting an exciting new avenue for statistical analysis of neuroimaging
783 data. Next steps will be to investigate how to interpret and report these posterior maps,
784 and to try more complex models that include additional model terms.

785 **Acknowledgements**

786 Computation has been performed on the HPC for Research cluster of the Berlin
787 Institute of Health.

788 **4.14 MOSAIC for VASO fMRI**

789 *Renzo (Laurentius) Huber, Remi Gau, Rüdiger Stirnberg, Philipp Ehses, Ömer Faruk Gülbán, Benedikt A. Poser*

790

791 Vascular Space Occupancy (VASO) is a functional magnetic resonance imaging (fMRI)
792 method that is used for high-resolution cortical layer-specific imaging⁶⁶. Currently, the
793 most popular sequence for VASO at modern SIEMENS scanners is the one by Stirnberg
794 and Stöcker⁶⁷ from the DZNE in Bonn, which is employed at more than 30 research labs
795 worldwide. This sequence concomitantly acquires fMRI BOLD and blood volume signals.
796 In the SIEMENS' reconstruction pipeline, these two complementary fMRI contrasts are
797 mixed together within the same time series, making the outputs counter-intuitive for
798 users. Specifically:

- 799 • The ‘raw’ NIfTI converted time-series are not BIDS compatible (see <https://github.com/bids-standard/bids-specification/issues/1001>).
800
- 801 • The order of odd and even BOLD and VASO image TRs is unprincipled, making
802 the ordering dependent on the specific implementation of NIfTI converters.

803 Workarounds with 3D distortion correction, results in interpolation artifacts. Altern-
804 ative workarounds without MOSAIC decorators result in unnecessarily large data sizes.

805 In the previous Brainhack¹, we extended the existing 3D-MOSAIC functor that was
806 previously developed by Benedikt Poser and Philipp Ehses. This functor had been previ-
807 ously used to sort volumes of images by dimensions of echo-times, by RF-channels, and by
808 magnitude and phase signals. In this Brainhack, we successfully extended and validated
809 this functor to also support the dimensionality of SETs (that is representing BOLD and
810 VASO contrast).

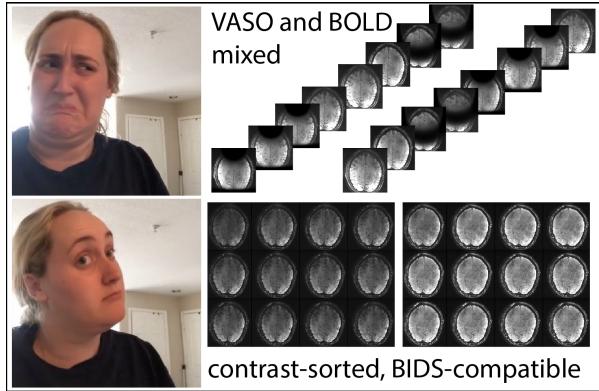


Figure 8: Previously, most VASO sequences provided unsorted image series of MRI contrasts. This was not BIDS compatible and could suffer from gradient non-linearity artifacts in the scanner’s MR-reconstruction pipeline. In Brainhack 2022, we adapted the SIEMENS reconstruction and to sort volume series by fMRI contrasts. This is BIDS compatible and does not require non-linearity corrections.

811 We are happy to share the compiled SIEMENS ICE (Image Calculation Environment)

812 functor that does this sorting. Current VASO users, who want to upgrade their recon-
813 struction pipeline to get the MOSAIC sorting feature too, can reach out to Renzo Huber
814 (RenzoHuber@gmail.com) or Rüdiger Stirnberg (Ruediger.Stirnberg@dzne.de).

815 Furthermore, Remi Gau, generated a template dataset that exemplifies how one could
816 to store layer-fMRI VASO data. This includes all the meta data for ‘raw and ‘derivatives’.
817 Link to this VASO fMRI BIDS demo: [https://gin.g-node.org/RemiGau/ds003216/
818 src/bids_demo](https://gin.g-node.org/RemiGau/ds003216/src/bids_demo).

819 Acknowledgements: We thank Chris Rodgers for instructions on how to overwrite
820 existing reconstruction binaries on the SIEMENS scanner without rebooting. We thank
821 David Feinberg, Alex Beckett and Samantha Ma for helping in testing the new reconstruc-
822 tion binaries at the Feinbergtron scanner in Berkeley via remote scanning. We thank
823 Maastricht University Faculty of Psychology and Neuroscience for supporting this project
824 with 2.5 hours of ‘development scan time’.

825 5 Conclusion and future directions

826 *Stefano Moia, Hao-Ting Wang*

827

828 Approaching the organisation of an event as an experiment allows incredible freedom
829 and dynamicity, albeit knowing that there will be risks and venues of improvement for
830 the future.

831 The organisation managed to provide a positive onsite environment, aiming to allow
832 participants to self-organise in the spirit of the Brainhack¹, with plenty of moral - and
833 physical - support.

834 The technical setup, based on heavy automatisation flow to allow project submission
835 to be streamlined, was a fundamental help to the organisation team, that would benefit
836 even more from the improvement of such automatisation flows.

837 This year, representatives of AFNI, FSL, and SPM (among the major neuroscience
838 software developers) took part in the event, and their presence was appreciated both by
839 other participants and themselves. In the future, connecting to more developers, not only
840 from the MRI community, might improve the quality of the Brainhack even more.

841 The most challenging element of the organisation was setting up an hybrid event.
842 While this element did not go as smoothly as it could have, this experimental setup seemed
843 to have worked, allowing the participation of about 70 participants online. However, there
844 is still a lot to improve for a truly hybrid event. For instance, it is important to allow spaces
845 (both in time and space) for participants on-site to interact with online participants, and
846 more attention, time, volunteers, and equipment has to be put to achieve a smooth
847 online participation. For this reason, the Open Science Special Interest Group instituted
848 a position to have a dedicated person for the hybridisation process. The other challenge
849 was to welcome newcomers into this heavily project-development-oriented event. While
850 newcomers managed to collaborate with projects and self-organise to learn open science
851 related skills, this integration of pre-event train track and beginner friendly process will
852 benefit from more attention.

853 Overall this OHBM Brainhack was a successful outcome for the organisation team
854 experiment, and we hope that our findings will be helpful to future Brainhack events
855 organisations.

856 **References**

- 857 1. Gau R, Noble S, Heuer K et al. Brainhack: Developing a culture of open, inclusive,
858 community-driven neuroscience. *Neuron*. 2021; 109(11):1769–1775.
- 859 2. Alkemade A, Bazin PL, Balesar R et al. A unified 3D map of microscopic architecture
860 and MRI of the human brain. *Science Advances*. 2022; 8(17):1–10.
- 861 3. Heuer K, Gulban OF, Bazin PL et al. Evolution of neocortical folding: A phylo-
862 genetic comparative analysis of MRI from 34 primate species. en. *Cortex*. 2019;
863 118():275–291.
- 864 4. Huntenburg JM, Steele CJ, Bazin PL. Nighres: processing tools for high-resolution
865 neuroimaging. en. *GigaScience*. 2018; 7(7).
- 866 5. Yushkevich PA, Piven J, Hazlett HC et al. User-guided 3D active contour seg-
867 mentation of anatomical structures: Significantly improved efficiency and reliability.
868 *NeuroImage*. 2006; 31(3):1116–1128.
- 869 6. Cox RW. *AFNI: Software for Analysis and Visualization of Functional Magnetic*
870 *Resonance Neuroimages*. Tech. rep. 1996:162–173.
- 871 7. Smith SM, Jenkinson M, Woolrich MW et al. Advances in functional and structural
872 MR image analysis and implementation as FSL. *NeuroImage*. Vol. 23. SUPPL. 1.
873 2004.
- 874 8. Gulban OF, Schneider M, Marquardt I, Haast RA, De Martino F. *A scalable method*
875 *to improve gray matter segmentation at ultra high field MRI*. Vol. 13. 6. 2018:1–31.
- 876 9. Huber L(R, Poser BA, Bandettini PA et al. LayNii: A software suite for layer-fMRI.
877 *NeuroImage*. 2021; 237(May):118091.
- 878 10. Apon A, Ngo L, Payne M, Wilson P. Assessing the effect of high performance
879 computing capabilities on academic research output. *Empirical Economics*. 2014;
880 48().
- 881 11. Oracle cloud sustainability. <https://www.oracle.com/uk/sustainability/green-cloud/>. Accessed: 2022-07-19.

- 883 12. *Oracle for Research*. <https://docs.oracle.com/en/programs/research/>. Ac-
884 cessed: 2022-07-19.
- 885 13. *NeuroDesk*. <https://www.neurodesk.org/>. Accessed: 2022-07-19.
- 886 14. Wilkinson MD, Dumontier M, Aalbersberg IJ et al. The FAIR Guiding Principles
887 for scientific data management and stewardship. en. *Sci. Data.* 2016; 3(1):160018.
- 888 15. Halchenko YO, Meyer K, Poldrack B et al. DataLad: distributed system for joint
889 management of code, data, and their relationship. *Journal of Open Source Software*.
890 2021; 6(63):3262.
- 891 16. King G. *An introduction to the dataverse network as an infrastructure for data*
892 *sharing*. 2007.
- 893 17. Hanke M, Pestilli F, Wagner AS, Markiewicz CJ, Poline JB, Halchenko YO. In
894 defense of decentralized research data management. en. *Neuroforum*. 2021; 27(1).
895 Publisher: De Gruyter Section: Neuroforum:17–25.
- 896 18. Hanke M, Poldrack B, Wagner AS et al. *datalad/datalad-osf: Cleanup*. Version 0.2.3.
897 Mar. 2021.
- 898 19. Halchenko YO, Hanke M, Heunis S et al. *DataLad-next extension*.
- 899 20. Jiang C, Selle A, Teran J. The Affine Particle-In-Cell Method Chenfanfu. *ACM*
900 *Transactions on Graphics*. 1965; 34(4):51.
- 901 21. Love E, Sulsky DL. An unconditionally stable, energy-momentum consistent imple-
902 mentation of the material-point method. *Computer Methods in Applied Mechanics*
903 *and Engineering*. 2006; 195(33-36):3903–3925.
- 904 22. Stomakhin A, Schroeder C, Chai L, Teran J, Selle A. A material point method for
905 snow simulation. *ACM Transactions on Graphics*. 2013; 32(4).
- 906 23. Ferrante O, Liu L, Minarik T et al. FLUX: A pipeline for MEG analysis. *NeuroIm-*
907 *age*. 2022; 253():119047.
- 908 24. Gramfort A, Luessi M, Larson E et al. MNE software for processing MEG and EEG
909 data. *NeuroImage*. 2014; 86():446–460.

- 910 25. Oostenveld R, Fries P, Maris E, Schoffelen JM. FieldTrip: Open source software for
911 advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput.*
912 *Intell. Neurosci.* 2011; 2011().
- 913 26. Appelhoff S, Sanderson M, Brooks TL et al. MNE-BIDS: Organizing electrophysiological
914 data into the BIDS format and facilitating their analysis. *Journal of Open*
915 *Source Software.* 2019; 4(44).
- 916 27. Carp J. On the plurality of (methodological) worlds: estimating the analytic flexibil-
917 ity of fMRI experiments. *Frontiers in neuroscience.* 2012; 6():149.
- 918 28. Kennedy DN, Abraham SA, Bates JF et al. Everything matters: the ReproNim
919 perspective on reproducible neuroimaging. *Frontiers in neuroinformatics.* 2019():1.
- 920 29. Schuff N, Woerner N, Boreta L et al. MRI of hippocampal volume loss in early
921 Alzheimer's disease in relation to ApoE genotype and biomarkers. *Brain.* 2009;
922 132(4):1067–1077.
- 923 30. Patenaude B, Smith SM, Kennedy DN, Jenkinson M. A Bayesian model of shape
924 and appearance for subcortical brain segmentation. *Neuroimage.* 2011; 56(3):907–
925 922.
- 926 31. Fischl B. FreeSurfer. *Neuroimage.* 2012; 62(2):774–781.
- 927 32. Xie L, Wisse LE, Pluta J et al. Automated segmentation of medial temporal lobe
928 subregions on in vivo T1-weighted MRI in early stages of Alzheimer's disease. *Human*
929 *brain mapping.* 2019; 40(12):3431–3451.
- 930 33. Tremblay-mercier J, Madjar C, Das S et al. Open science datasets from PREVENT-
931 AD, a longitudinal cohort of pre-symptomatic Alzheimer's disease. *NeuroImage:*
932 *Clinical.* 2021; 31():102733.
- 933 34. Esteban O, Birman D, Schaer M, Koyejo OO, Poldrack RA, Gorgolewski KJ. MRIQC:
934 Advancing the automatic prediction of image quality in MRI from unseen sites. *PloS*
935 *one.* 2017; 12(9):e0184661.
- 936 35. *MRIQC Documentation.* 2020. URL: <https://mriqc.readthedocs.io/en/stable/iqms/t1w.html>.

- 938 36. Urchs S, Dai A, Armoza J, Jahanpour A, Bhagwat N, Poline JB. *neurobagelapi*.
939 Version 0.2.0.
- 940 37. Markiewicz CJ, Gorgolewski KJ, Feingold F et al. The OpenNeuro resource for
941 sharing of neuroscience data. en. *Elife*. 2021; 10().
- 942 38. Gorgolewski KJ, Auer T, Calhoun VD et al. The brain imaging data structure, a
943 format for organizing and describing outputs of neuroimaging experiments. en. *Sci
944 Data*. 2016; 3():160044.
- 945 39. Field D, Garrity G, Gray T et al. The minimum information about a genome se-
946 quence (MIGS) specification. en. *Nat. Biotechnol.* 2008; 26(5):541–547.
- 947 40. Stang PE, Ryan PB, Racoosin JA et al. Advancing the science for active surveillance:
948 rationale and design for the Observational Medical Outcomes Partnership. en. *Ann.
949 Intern. Med.* 2010; 153(9):600–606.
- 950 41. Botvinik-nezer R, Holzmeister F, Camerer CF et al. Variability in the analysis of a
951 single neuroimaging dataset by many teams. en. 2020():26.
- 952 42. Nichols TE, Das S, Eickhoff SB et al. Best practices in data analysis and sharing in
953 neuroimaging using MRI. en. *Nature Neuroscience*. 2017; 20(3):299–303.
- 954 43. Gorgolewski KJ, Varoquaux G, Rivera G et al. NeuroVault.org: a web-based repos-
955 itory for collecting and sharing unthresholded statistical maps of the human brain.
956 en. *Frontiers in Neuroinformatics*. 2015; 9().
- 957 44. Mourik T van. *GiraffeToolBox*. Accessed: 2022-06-16.
- 958 45. Gorgolewski K. Nipype: a flexible, lightweight and extensible neuroimaging data
959 processing framework in Python. en. *Frontiers in Neuroinformatics*. 2017():15.
- 960 46. Poldrack R, Kittur A, Kalar D et al. The Cognitive Atlas: Toward a Knowledge
961 Foundation for Cognitive Neuroscience. *Frontiers in Neuroinformatics*. 2011; 5().
- 962 47. Dockès J, Poldrack RA, Primet R et al. NeuroQuery, comprehensive meta-analysis
963 of human brain mapping. *eLife*. 2020; 9(). Ed. by Büchel C, Yeo T, Wager TD.
964 Publisher: eLife Sciences Publications, Ltd:e53385.

- 965 48. Yarkoni T, Poldrack RA, Nichols TE, Van essen DC, Wager TD. Large-scale auto-
966 mated synthesis of human functional neuroimaging data. en. *Nat Methods*. 2011;
967 8(8). Number: 8 Publisher: Nature Publishing Group:665–670.
- 968 49. Weber MJ, Thompson-schill SL. Functional neuroimaging can support causal claims
969 about brain function. *J Cogn Neurosci*. 2010; 22(11):10.1162/jocn.2010.21461.
- 970 50. Siddiqi SH, Kording KP, Parvizi J, Fox MD. Causal mapping of human brain func-
971 tion. en. *Nat Rev Neurosci*. 2022; 23(6). Number: 6 Publisher: Nature Publishing
972 Group:361–375.
- 973 51. Price CJ. The evolution of cognitive models: From neuropsychology to neuroimaging
974 and back. en. *Cortex*. In Memory of Professor Glyn Humphreys 2018; 107():37–49.
- 975 52. Larivière S, Paquola C, Park By et al. The ENIGMA Toolbox: multiscale neural
976 contextualization of multisite neuroimaging datasets. en. *Nat Methods*. 2021; 18(7).
977 Number: 7 Publisher: Nature Publishing Group:698–700.
- 978 53. Chen JE, Lewis LD, Chang C et al. Resting-State “Physiological Networks”. *NeuroIm-*
979 *age*. 2020; 213():116707.
- 980 54. Bulte D, Wartolowska K. Monitoring Cardiac and Respiratory Physiology during
981 fMRI. *NeuroImage*. 2017; 154():81–91.
- 982 55. Caballero-gaudes C, Reynolds RC. Methods for Cleaning the BOLD fMRI Signal.
983 *NeuroImage*. 2017; 154():128–149.
- 984 56. Alcalá D, Ayyagari A, Bottenhorn K et al. *physiopy/phys2bids: BIDS formatting of*
985 *physiological recordings*. June 2021.
- 986 57. Chen G, Xiao Y, Taylor PA et al. Handling Multiplicity in Neuroimaging through
987 Bayesian Lenses with Multilevel Modeling. *Neuroinformatics*. 2019; 17(4):515–545.
- 988 58. Chen G, Taylor PA, Stoddard J, Cox RW, Bandettini PA, Pessoa L. Sources of
989 Information Waste in Neuroimaging: Mishandling Structures, Thinking Dichotom-
990 ously, and Over-Reducing Data. *Aperture Neuro*. 2022; 2021(5):46.

- 991 59. Bürkner PC. brms: An R Package for Bayesian Multilevel Models Using Stan. en.
992 *Journal of Statistical Software*. 2017; 80(1). Number: 1:1–28.
- 993 60. Salvatier J, Wiecki TV, Fonnesbeck C. Probabilistic programming in Python using
994 PyMC3. *PeerJ Computer Science*. 2016; 2():e55.
- 995 61. Capretto T, Piho C, Kumar R, Westfall J, Yarkoni T, Martin OA. *Bambi: A simple*
996 *interface for fitting Bayesian linear models in Python*. 2020. arXiv: 2012 . 10754
997 [stat.CO].
- 998 62. Wakeman DG, Henson RN. A multi-subject, multi-modal human neuroimaging
999 dataset. en. *Scientific Data*. 2015; 2(1):150001.
- 1000 63. Waller L, Erk S, Pozzi E et al. ENIGMA HALFpipe: Interactive, reproducible,
1001 and efficient analysis for resting-state and task-based fMRI data. en. *Human Brain*
1002 *Mapping*. 2022; 43(9):2727–2742.
- 1003 64. Esteban O, Markiewicz CJ, Blair RW et al. fMRIPrep: a robust preprocessing
1004 pipeline for functional MRI. en. *Nature Methods*. 2019; 16(1). Number: 1 Publisher:
1005 Nature Publishing Group:111–116.
- 1006 65. Phan D, Pradhan N, Jankowiak M. *Composable Effects for Flexible and Accelerated*
1007 *Probabilistic Programming in NumPyro*. arXiv:1912.11554 [cs, stat]. Dec. 2019.
- 1008 66. Huber L, Finn ES, Chai Y et al. Layer-dependent functional connectivity methods.
1009 *Progress in Neurobiology*. 2021; 207().
- 1010 67. Stirnberg R, Stöcker T. Segmented K-Space Blipped-Controlled Aliasing in Parallel
1011 Imaging (Skipped-CAIPI) for High Spatiotemporal Resolution Echo Planar Imaging.
1012 *Magnetic Resonance in Medicine*. 2021; 85(0):1540–1551.