The future of data analysis is now: Integrating generative AI in neuroimaging methods development

Elizabeth DuPre[a], Russell Alan Poldrack[a]

[a]Department of Psychology, Stanford University, Stanford, CA, United States

Corresponding Author: Elizabeth DuPre (elizabethm.dupre@gmail.com)

ABSTRACT

In this perspective, we highlight how emerging artificial intelligence tools are likely to impact the experiences of researchers conducting computational fMRI analyses. While calls for the automatization of statistical procedures date back at least to the inception of "data science" as a field, generative artificial intelligence offers new opportunities to advance field practice. We highlight how these tools are poised to impact both new neuroimaging methods development in areas such as image quality control and in day-to-day practice when generating analysis code. We argue that considering generative artificial intelligence as a catalyst for computational neuroscience—rather than as unique tools in their own right—can substantially improve its positioning in the research ecosystem. In particular, we argue that generative artificial intelligence will reinforce the importance of existing open science initiatives, rather than supplanting them. Overall, we call for clearer metrics by which neuroimaging results—whether generated by individual research teams or by generative artificial intelligence technologies—can be meaningfully compared.

Keywords: computational science, open science, neuroimaging methods, artificial intelligence, code assistants

1. INTRODUCTION

Alongside fields such as genomics and astronomy, neuroimaging has transformed over the last decade into a data intensive, "Big Data" field. With the emergence of funder-mandated data sharing (Jwa & Poldrack, 2022), individual researchers now have access to more neuroimaging data than ever before, with 406 terabytes accessed through the data sharing platform OpenNeuro from May 2020 to April 2021 alone (Markiewicz et al., 2021). While this growth has presented important opportunities to improve research practices (Gratton et al., 2022), it has also created unanticipated challenges. Researchers now need exponentially more training in data science and statistics than they did a decade ago. Many research supervisors do not have the necessary training to properly store, access, and analyze their data, and graduate degree programs lag behind evolving field standards. Open science communities such as Brainhack (Gau et al., 2021) as well as targeted educational initiatives such as NeuroHackademy (Huppenkothen et al., 2018) and NeuroMatch Academy (van Viegen et al., 2021) have stepped in to fill this gap with training materials and peer support networks. However, the scale of this challenge continues to outpace relatively slow changes both to training programs and to institutional hiring decisions, where relatively few roles focus on data science expertise.

In this space, the emergence of generative artificial intelligence (AI) has promised to revolutionize researchers' experiences of the scientific process. Generative artificial intelligence (generative AI) is a class of algorithms that combine generative models with deep neural networks, allowing users to generate probable new text,

images, and audio outputs from supplied human inputs such as text or image prompts. AI-assisted coding, in particular, has been argued to improve developer productivity—particularly in supporting relatively novice programmers (Peng et al., 2023; c.f. Orosz & Beck, 2023)—and has anecdotally been suggested to provide a more enjoyable coding experience by automating low-level details and reducing time to problem solution. Its potential to improve the scientific process is thus clear, though how it and other generative AI tooling are adapted by the broader community will ultimately shape its impact. There is thus a clear potential that it could be used to similarly accelerate methods development. While other reviews have considered emerging use cases of generative AI models for neuroimaging analyses (Gong et al., 2023), the potential impact of these methods on the scientific process itself remains underexplored. Here, we aim to directly consider these questions, overviewing the potential impact of AI-assisted tooling on neuroimaging field practice. We note that—while another important, related topic in this area—the influence of AI technologies on scientific writing and reviewing has ethical considerations for the current academic credit system (Nature, 2023), but we leave those for future discussion. Instead, we highlight how the emergence of generative AI methods is likely to impact both day-to-day practice and new analytic methods for neuroimaging. We further argue that these methods will play against the well-established open source ecosystem in neuroimaging, introducing new considerations for both developers and users of these tools.

2.   A HISTORICAL PERSPECTIVE: THE INTRODUCTION OF DATA SCIENCE

While the mass adoption of generative AI technologies is a unique historical moment, we can use other methodological turning points to frame these developments. Perhaps the most relevant of these is the creation of "data science" as an approach to systematizing scientific data analysis. In 1962, John Tukey published "The Future of Data Analysis" in which he called for the creation of data science—a field contrasted with statistics in its unique focus on data analysis (Tukey, 1962).

Among other prescient arguments, Tukey emphasized the importance of developing automated, standardized approaches to as many statistical procedures as possible rather than relying on individual researcher expertise. Failure to do so, he suggested, would result in poor research practices as more data became available, since "most data analysis is going to be done by people who are not sophisticated data analysts and who have very limited time; if you do not provide them tools the data will

be even less studied" (p. 22). Indeed, neuroimaging has found itself in exactly this position, as a lack of dedicated data science training siloed analytic expertise both within and across labs. The emergence of data standards such as BIDS (Gorgolewski et al., 2016) has allowed the Nipreps ecosystem (Esteban et al., 2019) to fill this gap in part by automating some common fMRI preprocessing workflows based on existing software platforms such as AFNI (Cox, 1996), FSL (Jenkinson et al., 2012), and ANTS (Tustison et al., 2021).

While extensions to other neuroimaging modalities are an obvious next step (though see, e.g., MNE-BIDS Pipeline for one existing example; Appelhoff et al., 2019), work across the Nipreps ecosystem has highlighted two core areas of work for neuroimaging methods development. First, the need to continue standardizing analysis practices in areas that have yet to develop a formal theory, such as image quality control. Second, the need to facilitate automation in the analysis of preprocessed fMRI datasets, where the exact analysis choices are specific to a particular research question and task design. Generative artificial intelligence (AI) may help to address both of these broad challenges.

Neuroscience occupies a relatively unique position among data-intensive fields in that AI promises to significantly impact both its methodological and theoretical foundations through, for example, "NeuroAI" (Richards et al., 2019; Zador et al., 2023). Even when considering only its methodological impacts, AI has the potential to cause significant harm—delaying rather than accelerating field progress—if researchers are ill-informed about the capabilities of these tools. We first consider how generative AI is likely to impact one specific aspect of neuroimaging data processing: quality control. We then consider how the more recent development of large-language models such as ChatGPT and GitHub Copilot[1] is beginning to change neuroimaging method development, and how these developments may impact the broader tooling ecosystem.

3.   USING GENERATIVE AI FOR NEW METHODS IN IMAGE QUALITY CONTROL

While Tukey persuasively argued for automating existing statistical procedures, many neuroimaging methods do not have a clear path to automation. In particular, experimental procedures that do not have a clear metric of success are especially difficult to automate, as

---

[1] Although we list specific models here for clarity, many of the larger language and image generation models are currently facing legal challenges for their use of associated training data. The reader should thus be alert to changes in the legal status of these models.

different scientists often prioritize different goals. One example is in experimental stimulus generation: while generative AI tools such as MidJourney and StableDiffusion can easily create a huge variety of images, individual researchers must choose which to prioritize in limited experimental time.

An even clearer example is in image quality control, where visual inspection remains the gold standard for the field. Researchers develop unique expertise in image quality control based on their training, largely in the context of specific scientific questions. It is not surprising, then, that quality control ratings differ even between experts (Taylor et al., 2023). Despite this uncertainty, the sheer amount of neuroimaging data to be checked has pushed researchers toward developing no-reference image quality metrics, with the idea that these can guide researcher inspection and downstream machine learning analyses (Bastiani et al., 2019; Esteban et al., 2017). This idea of using AI and machine learning techniques to improve quality control is not unique to neuroscience. As data cleaning comprises a significant portion of analyst time in industry settings, automated quality control for tabular data is an active area of research. While projects such as DirtyData (https://project.inria.fr/dirtydata/; Perez-Lebel et al., 2022) aim to provide theoretical guarantees in data cleaning, existing work points toward relatively nuanced—rather than robust—applications, in which the details of the problem drive performance (Cheng et al., 2022).

Despite this challenge, there is reason for optimism. In neuroimaging preprocessing, projects such as NoBrainer (Kaczmarzyk et al., 2023) and FastSurfer (Henschel et al., 2020) have used AI to dramatically reduce compute time for image-based tasks such as brain segmentation while maintaining high-quality outputs. There is, therefore, an obvious potential in using AI methods in neuroimaging data processing. Critically, however, these existing tools document their extensive validation against a large number of publicly available, labeled datasets. To date, large, labeled datasets for quality control validation are much harder to acquire.

One potential solution is to augment existing labels using crowd sourcing. Efforts such as Swipes for Science (Keshavan et al., 2019) and Brainmatch (Benhajali et al., 2020) have successfully generated quality control labels at scale using distributed, citizen science efforts. While these approaches have shown significant potential (Richie-Halford et al., 2022), the increasing availability of consumer-grade AI calls future work on crowdsourced platforms into question. That is, an increasing number of crowd workers are using tools such as ChatGPT to complete online crowd-sourced tasks, particularly those that allow for free-text input (e.g., to describe artifacts; Veselovsky et al., 2023). Training an algorithm on these

AI-generated responses could lead to "model collapse," in which performance significantly degrades over time (Shumailov et al., 2023).

More recently, efforts such as the Demonstrating Quality Control Procedures project (https://osf.io/qaesm/) have turned to focus on the task of label generation itself. The resulting quality control protocols may in turn be critical for future AI applications; indeed, generative AI could assist in this effort, by augmenting available data with a specific label (e.g., a "ringing artifact") to assess inter-rater agreement. Certain quality control procedures and resulting labels, however, are likely to be specific to a given population. For example, lesion segmentation in stroke patients is an important image processing step that is not shared with normative population samples (Liew et al., 2018). In these cases, both limited data availability and strong privacy concerns are likely to impair extensive data labeling efforts.

Alternative learning paradigms such as federated learning provide one path forward, by allowing joint models to be learned via model sharing rather than data sharing (Stripelis et al., 2022). Similarly, neuroimaging foundation models (Caro et al., 2023; Thomas et al., 2022) may allow for transferring pretrained, data-intensive models to smaller, more specialized datasets. Importantly, however, both of these paradigms require a strong degree of data standardization to ensure that model features transfer appropriately between contexts, for example, experimental sites.

4. AI-ASSISTED CODING FOR WELL-DEFINED CONTEXTS

While Tukey encouraged the automatization of as many statistical procedures as possible, key neuroimaging analysis steps have resisted automation. This is in part because of the range of different analytic choices that a single dataset can support (Carp, 2012), making automation nontrivial. For example, two of the most widely used analyses of fMRI data are first- and second-level General Linear Modelling (GLM; Friston et al., 1995), which have served as cornerstones of fMRI analysis for over two decades. Only in the past few years, however, has dedicated tooling emerged to represent these very common analyses in a standardized format (e.g., FitLins; Markiewicz et al., 2022). Without a standardized and machine-readable input and output structure, automation through AI or other tooling is difficult.

As standardized GLM analyses are still being actively developed, it is no surprise that other, newer methods largely lack standardized formats. In these cases, it falls to the experimenter to ensure that their generated code is correctly executing the desired analysis. For researchers

---

BOX 1. EXAMPLE WORKFLOW FOR CODE GENERATION.

Working with large language models can accelerate computational analyses. While many workflows are possible when working with these models to generate code, we provide one generic example here to help orient the reader.

1. Use the developer API to set a random seed value and track the system fingerprint, whenever possible, to assist in reproducing the system behavior.
2. Prompt the model with a specific query. Some suggestions[2] include:
    - "Create Python code to generate new test data for the following function."
    - "Create a Python class to perform spatial smoothing on a NifTI image using a Gaussian filter."
    - "Given the following function, generate associated tests."
    - "Add docstrings to the included code."
    - "Refactor the provided code to make it easier to read."
3. Vary the random seed and temperature to better understand the variability of behavior of the model across parameterizations and adapt suggested code accordingly.
4. Test that the model-suggested code runs on real data and generates sensible outputs, if applicable.

---

with no or minimal training in software engineering, this requirement can block access to certain scientific questions—or yield incorrect inferences (Soergel, 2014). In these cases, AI-assisted coding is an appealing solution to act as a "catalytic" research accelerator, lowering the activation energy of producing high-quality code. It does not provide a full free lunch, however. Instead, it shifts the burden from generating code to reviewing and testing code (Poldrack et al., 2023). For example, if AI-assisted coding is used to generate tests, does a failing test indicate a problem with the test or the tested code? Does the derived code produce sensible inputs and outputs? These questions still require active input from researchers who are comfortable with software engineering principles.

While this approach is tractable at some levels of analysis (e.g., to write a single function within a larger workflow), scaling it to a complete neuroimaging experiment exponentially increases the possibility of errors. Since there is rarely a single metric which researchers can benchmark their results against, it is also difficult to isolate when AI-generated code is producing meaningful variability or introducing subtle errors. Having access to data and code supporting previous results can provide a meaningful sanity check, serving as a higher-level test. That is, researchers can assess: With the assistance of GitHub Copilot or ChatGPT, can I generate a version of this existing code that runs on my data? If I re-run it on the original data, do I see similar results to the original analysis? This provides an index of how robust and replicable (The Turing Way Community, 2023) a previous result is, as well as a confirmation that new

code is correctly reproducing the intended analyses. We provide an example workflow for generating new research code using these tools in Box 1. Critically, however, it relies on the availability of open data and code for existing experiments.

The importance of data and code sharing in this context points toward open science solutions. We thus turn next to the interaction of AI-driven science with the existing open source and open science landscapes.

5. OPEN DATA, OPEN CODE, OPEN AI? HOW AI-ASSISTED TOOLING WILL INTERACT WITH OPEN SCIENCE

Open science has been a critical driver of methods development in neuroimaging over the last decade, yielding large-scale field standardization efforts for the most common preprocessing workflows (Kiar et al., 2023). Open science was not a specific focus of "The Future of Data Analysis" (Tukey, 1962); this reflects the change in the research landscape, however, rather than the potential effects of open science efforts. Indeed, in the years since "The Future of Data Analysis" was published, open science practices of data and code sharing have become a driving force both in neuroimaging and in many other data intensive fields (Donoho, 2017). As reviewed above, it is clear that many of these same open science practices interact with AI-assisted methods. In particular, AI-assisted coding will catalyze new research by lowering the barrier for coding, allowing researchers with little formal data science training to run otherwise inaccessible neuroimaging analyses. Additional generative AI technologies are likely to be useful in refining fMRI methods, such as neuroimaging quality control.

---

[2] Adapted from Poldrack et al. (2023).

An obvious question is how AI-assisted tooling will change open science and its practice in neuroimaging methods development. That is, will AI-assisted tooling diminish open science practices such as data and code sharing, when data and code can simply be created on command? Evidence from other fields suggests the opposite. Recently, David Donoho—a leading figure in the field of data science—argued that the commercial success of AI reflects a robust data science culture in empirical machine learning (Donoho, 2023). He advocated for not only open code and data sharing, but also explicit metrics by which individual analyses can be compared, such as performance on a public prediction challenge. While prediction challenges have seen limited success in neuroimaging to date (Traut et al., 2022), the idea of clear metrics by which neuroimaging results can be compared will be critically important to more fully adopt generative AI technologies into neuroimaging methods development. Without clear outcome metrics, strong human-in-the-loop systems will be necessary to review individual AI applications, falling short of Tukey's call for automated analysis. Greater standardization efforts are, therefore, a critical step in adopting generative AI technologies. Encouragingly, open science efforts such as the Demonstrating Quality Control Procedures project are beginning to chart a path toward standardization for previously ambiguous metrics, such as visual quality control. Nonetheless, significant work remains to be done, including making the results of neuroimaging analyses FAIR (Findable, Accessible, Interoperable, and Reusable; Poline et al., 2022) so that they can be objectively assessed by independent investigators. While generative AI technologies may help in this process (e.g., by developing new examples for human labeling), deeper progress will remain tied to human-led standardization efforts.

Focusing on the idea of automated analysis suggests that generative AI will not replace existing open science initiatives any time soon. Instead, it will require researchers to produce not only open data and open code, but also clear outcomes by which individual experiments can be compared. Initiatives such as the Neuroimaging Analysis Replication and Prediction Study (NARPS; Botvinik-Nezer et al., 2020) confirm that teams of researchers conducting the same experiments on the same neuroimaging data can produce substantial differences in results. This work does not provide, however, a clear framework to compare each team's results. Multiverse analyses (Dafflon et al., 2022) or more general "vibration" analyses (Bhagwat et al., 2021) may be a useful mechanism to calibrate the range of possible results for a given experiment. Ideally, these calibrations could guide future work to generate open metrics. For the moment, however, continuing to openly share code and data to allow for these analyses—either within or across research teams—will be necessary to better situate emerging AI-assisted methods.

6. CONCLUSIONS

As a data intensive field, neuroimaging methods are increasingly reliant on data science practices to drive methodological innovation. Nonetheless, most researchers lack necessary data science training. Generative AI tooling may help to fill this gap; in doing so, however, it will strongly interact with existing systems for neuroimaging methods development. Many of these have matured in coordination with open science efforts, particularly those that focus on data and code sharing. In our view, the emergence of AI tooling will not supplant these efforts but instead reinforce their importance.

That is not to say that the interaction between open science and generative AI methods will be entirely synergistic. For example, defaced neuroimaging data are considered private data under The European Union's General Data Protection Regulations (GDPR), and nonanonymized neuroimaging data are considered private by even more countries, such as the United States. Directly providing neuroimages to proprietary tools such as ChatGPT or GitHub Copilot, therefore, is broadly unethical. Other examples of potential AI and open science interactions are somewhat murkier. For example, AI-assisted coding could conceivably lower the barrier to collaboration with Research Software Engineers (RSEs; Woolston, 2022) or with the broader open source ecosystem of neuroimaging tools, such as NIPY (Brett et al., 2009). That is, by making it easier to produce code, researchers may be able to more easily contribute to open source tools, supporting more analytic choices than would otherwise be possible. Nonetheless, there are several potential pitfalls to this approach, especially given the already limited developer time for code review and maintenance in open source projects (Eghbal, 2016) and the limited institutional support for RSEs to date.

Overall, we suggest that best capitalizing on AI-assisted tooling will require returning to the core principles of data science. In particular, active work on clear metrics by which research results can be compared is likely to significantly improve both the adoption of AI-assisted tooling and the state of neuroimaging methods as a whole.

DATA AND CODE AVAILABILITY

No additional datasets or code are associated with this paper.

AUTHOR CONTRIBUTIONS

DECLARATION OF COMPETING INTEREST

ACKNOWLEDGEMENTS

REFERENCES

Appelhoff, S., Sanderson, M., Brooks, T. L., van Vliet, M., Quentin, R., Holdgraf, C., Chaumon, M., Mikulan, E., Tavabi, K., Höchenberger, R., Welke, D., Brunner, C., Rockhill, A. P., Larson, E., Gramfort, A., & Jas, M. (2019). MNE-BIDS: Organizing electrophysiological data into the BIDS format and facilitating their analysis. *Journal of Open Source Software*, *4*(44), 1896. https://doi.org/10.21105/joss.01896

Bastiani, M., Cottaar, M., Fitzgibbon, S. P., Suri, S., Alfaro-Almagro, F., Sotiropoulos, S. N., Jbabdi, S., & Andersson, J. L. R. (2019). Automated quality control for within and between studies diffusion MRI data using a non-parametric framework for movement and distortion correction. *NeuroImage*, *184*, 801–812. https://doi.org/10.1016/j.neuroimage.2018.09.073

Benhajali, Y., Badhwar, A., Spiers, H., Urchs, S., Armoza, J., Ong, T., Pérusse, D., & Bellec, P. (2020). A standardized protocol for efficient and reliable quality control of brain registration in functional MRI studies. *Frontiers in Neuroinformatics*, *14*, 7. https://doi.org/10.3389/fninf.2020.00007

Bhagwat, N., Barry, A., Dickie, E. W., Brown, S. T., Devenyi, G. A., Hatano, K., DuPre, E., Dagher, A., Chakravarty, M., Greenwood, C. M. T., Misic, B., Kennedy, D. N., & Poline, J.-B. (2021). Understanding the impact of preprocessing pipelines on neuroimaging cortical surface analyses. *Gigascience*, *10*(1), giaa155. https://doi.org/10.1093/gigascience/giaa155

Botvinik-Nezer, R., Holzmeister, F., Camerer, C. F., Dreber, A., Huber, J., Johannesson, M., Kirchler, M., Iwanir, R., Mumford, J. A., Adcock, R. A., Avesani, P., Baczkowski, B. M., Bajracharya, A., Bakst, L., Ball, S., Barilari, M., Bault, N., Beaton, D., Beitner, J., … Schonberg, T. (2020). Variability in the analysis of a single neuroimaging dataset by many teams. *Nature*, *582*(7810), 84–88. https://doi.org/10.1038/s41597-019-0113-7

Brett, M., Taylor, J., Burns, C., Millman, K. J., Perez, F., Roche, A., Thirion, B., & D'Esposito, M. (2009). NIPY: An open library and development framework for FMRI data analysis. *NeuroImage*, *47*(Suppl 1), S196. https://doi.org/10.1016/s1053-8119(09)72223-2

Caro, J. O., de O. Fonseca, A. H., Averill, C., Rizvi, S. A., Rosati, M., Cross, J. L., Mittal, P., Zappala, E., Levine, D., Dhodapkar, R. M., Abdallah, C. G., & van Dijk, D. (2023). BrainLM: A foundation model for brain activity recordings. *bioRxiv*, 2023.09.12.557460. https://doi.org/10.1101/2023.09.12.557460

Carp, J. (2012). On the plurality of (methodological) worlds: Estimating the analytic flexibility of fMRI experiments. *Frontiers in Neuroscience*, *6*, 149. https://doi.org/10.3389/fnins.2012.00149

Cheng, C., Asi, H., & Duchi, J. (2022). How many labelers do you have? A closer look at gold-standard labels. *arXiv:2206.12041*. https://arxiv.org/abs/2206.12041

Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, *29*(3), 162–173. https://doi.org/10.1006/cbmr.1996.0014

Dafflon, J., F Da Costa, P., Váša, F., Monti, R. P., Bzdok, D., Hellyer, P. J., Turkheimer, F., Smallwood, J., Jones, E., & Leech, R. (2022). A guided multiverse study of neuroimaging analyses. *Nature Communications*, *13*(1), 3758. https://doi.org/10.1038/s41467-022-31347-8

Donoho, D. (2017). 50 years of data science. *Journal of Computational and Graphical Statistics*, *26*(4), 745–766. https://doi.org/10.1080/10618600.2017.1384734

Donoho, D. (2023). Data science at the singularity. *arXiv:2310.00865*. https://arxiv.org/abs/2310.00865

Eghbal, N. (2016). *The unseen labor behind our digital infrastructure* (Rep.). Ford Foundation. https://www.fordfoundation.org/work/learning/research-reports/roads-and-bridges-the-unseen-labor-behind-our-digital-infrastructure/

Esteban, O., Birman, D., Schaer, M., Koyejo, O. O., Poldrack, R. A., & Gorgolewski, K. J. (2017). MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites. *PLoS One*, *12*(9), e0184661. https://doi.org/10.1371/journal.pone.0184661

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., & Gorgolewski, K. J. (2019). fMRIPrep: A robust preprocessing pipeline for functional MRI. *Nature Methods*, *16*(1), 111–116. https://doi.org/10.1038/s41592-018-0235-4

Friston, K. J., Holmes, A. P., Poline, J. B., Grasby, P. J., Williams, S. C., Frackowiak, R. S., & Turner, R. (1995). Analysis of fMRI time-series revisited. *NeuroImage*, *2*(1), 45–53. https://doi.org/10.1006/nimg.1995.1007

Gau, R., Noble, S., Heuer, K., Bottenhorn, K. L., Bilgin, I. P., Yang, Y.-F., Huntenburg, J. M., Bayer, J. M. M., Bethlehem, R. A. I., Rhoads, S. A., Vogelbacher, C., Borghesani, V., Levitis, E., Wang, H.-T., Van Den Bossche, S., Kobeleva, X., Legarreta, J. H., Guay, S., Atay, S. M., … Brainhack Community. (2021). Brainhack: Developing a culture of open, inclusive, community-driven neuroscience. *Neuron*, *109*(11), 1769–1775. https://doi.org/10.1016/j.neuron.2021.04.001

Gong, C., Jing, C., Chen, X., Pun, C. M., Huang, G., Saha, A., Nieuwoudt, M., Li, H.-X., Hu, Y., & Wang, S. (2023). Generative AI for brain image computing and brain network computing: A review. *Frontiers in Neuroscience*, *17*, 1203104. https://doi.org/10.3389/fnins.2023.1203104

Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., Flandin, G., Ghosh, S. S., Glatard, T., Halchenko, Y. O., Handwerker, D. A., Hanke, M., Keator, D., Li, X., Michael, Z., Maumet, C., Nichols, B. N., Nichols, T. E., Pellman, J., … Poldrack, R. A. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, *3*, 160044. https://doi.org/10.1038/sdata.2016.44

Gratton, C., Nelson, S. M., & Gordon, E. M. (2022). Brain-behavior correlations: Two paths toward reliability. *Neuron*, 110(9), 1446–1449. https://doi.org/10.1016/j.neuron.2022.04.018

Henschel, L., Conjeti, S., Estrada, S., Diers, K., Fischl, B., & Reuter, M. (2020). FastSurfer—A fast and accurate deep learning based neuroimaging pipeline. *NeuroImage*, 219, 117012. https://doi.org/10.1016/j.neuroimage.2020.117012

Huppenkothen, D., Arendt, A., Hogg, D. W., Ram, K., VanderPlas, J. T., & Rokem, A. (2018). Hack weeks as a model for data science education and collaboration. *Proceedings of the National Academy of Sciences of the United States of America*, 115(36), 8872–8877. https://doi.org/10.1073/pnas.1717196115

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, 62(2), 782–790. https://doi.org/10.1016/j.neuroimage.2011.09.015

Jwa, A. S., & Poldrack, R. A. (2022). The spectrum of data sharing policies in neuroimaging data repositories. *Human Brain Mapping*, 43(8), 2707–2721. https://doi.org/10.1002/hbm.25803

Kaczmarzyk, J., McClure, P., Zulfikar, W., Rana, A., Rajaei, H., Richie-Halford, A., Bansal, S., Jarecka, D., Lee, J., & Ghosh, S. (2023). Neuronets/nobrainer: 1.0.0. https://doi.org/10.5281/zenodo.4995077

Keshavan, A., Yeatman, J. D., & Rokem, A. (2019). Combining citizen science and deep learning to amplify expertise in neuroimaging. *Frontiers in Neuroinformatics*, 13, 29. https://doi.org/10.3389/fninf.2019.00029

Kiar, G., Clucas, J., Feczko, E., Goncalves, M., Jarecka, D., Markiewicz, C. J., Halchenko, Y. O., Hermosillo, R., Li, X., Miranda-Dominguez, O., Ghosh, S., Poldrack, R. A., Satterthwaite, T. D., Milham, M. P., & Fair, D. (2023). Align with the NMIND consortium for better neuroimaging. *Nature Human Behavior*, 7(7), 1027–1028. https://doi.org/10.1038/s41562-023-01647-0

Liew, S.-L., Anglin, J. M., Banks, N. W., Sondag, M., Ito, K. L., Kim, H., Chan, J., Ito, J., Jung, C., Khoshab, N., Lefebvre, S., Nakamura, W., Saldana, D., Schmiesing, A., Tran, C., Vo, D., Ard, T., Heydari, P., Kim, B., … Stroud, A. (2018). A large, open source dataset of stroke anatomical brain images and manual lesion segmentations. *Scientific Data*, 5, 180011. https://doi.org/10.1038/sdata.2018.11

Markiewicz, C. J., De La Vega, A., Wagner, A., Halchenko, Y. O., Finc, K., Ciric, R., Goncalves, M., Nielson, D. M., Kent, J. D., Lee, J. A., Bansal, S., Poldrack, R. A., & Gorgolewski, K. J. (2022). Poldracklab/fitlins: 0.11.0. https://doi.org/10.5281/zenodo.7217447

Markiewicz, C. J., Gorgolewski, K. J., Feingold, F., Blair, R., Halchenko, Y. O., Miller, E., Hardcastle, N., Wexler, J., Esteban, O., Goncalves, M., Jwa, A., & Poldrack, R. (2021). The OpenNeuro resource for sharing of neuroscience data. *Elife*, 10, e71774. https://doi.org/10.7554/elife.71774

Nature. (2023). Tools such as ChatGPT threaten transparent science; here are our ground rules for their use. *Nature*, 613(7945), 612. https://doi.org/10.1038/d41586-023-00191-1

Orosz, G., & Beck, K. (2023). *Measuring developer productivity? A response to McKinsey*. Retrieved October 25, 2023, from https://newsletter.pragmaticengineer.com/p/measuring-developer-productivity

Peng, S., Kalliamvakou, E., Cihon, P., & Demirer, M. (2023). The impact of AI on developer productivity: Evidence from GitHub copilot. *arXiv:2302.06590*. https://arxiv.org/abs/2302.06590

Perez-Lebel, A., Varoquaux, G., Le Morvan, M., Josse, J., & Poline, J.-B. (2022). Benchmarking missing-values approaches for predictive models on health databases. *Gigascience*, 11, giac013. https://doi.org/10.1093/gigascience/giac013

Poldrack, R. A., Lu, T., & Beguš, G. (2023). AI-assisted coding: Experiments with GPT-4. *arXiv:2304.13187*. https://arxiv.org/abs/2304.13187

Poline, J.-B., Kennedy, D. N., Sommer, F. T., Ascoli, G. A., Van Essen, D. C., Ferguson, A. R., Grethe, J. S., Hawrylycz, M. J., Thompson, P. M., Poldrack, R. A., et al. (2022). Is neuroscience FAIR? A call for collaborative standardisation of neuroscience data. *Neuroinformatics*, 20(2), 507–512. https://doi.org/10.1007/s12021-021-09557-0

Richards, B. A., Lillicrap, T. P., Beaudoin, P., Bengio, Y., Bogacz, R., Christensen, A., Clopath, C., Costa, R. P., de Berker, A., Ganguli, S., Gillon, C. J., Hafner, D., Kepecs, A., Kriegeskorte, N., Latham, P., Lindsay, G. W., Miller, K. D., Naud, R., Pack, C. C., … Kording, K. P. (2019). A deep learning framework for neuroscience. *Nature Neuroscience*, 22(11), 1761–1770. https://doi.org/10.1038/s41593-019-0520-2

Richie-Halford, A., Cieslak, M., Ai, L., Caffarra, S., Covitz, S., Franco, A. R., Karipidis, I. I., Kruper, J., Milham, M., Avelar-Pereira, B., Roy, E., Sydnor, V. J., Yeatman, J. D., Fibr Community Science Consortium, Satterthwaite, T. D., & Rokem, A. (2022). An analysis-ready and quality controlled resource for pediatric brain white-matter research. *Scientific Data*, 9(1), 616. https://doi.org/10.1038/s41597-022-01695-7

Shumailov, I., Shumaylov, Z., Zhao, Y., Gal, Y., Papernot, N., & Anderson, R. (2023). The curse of recursion: Training on generated data makes models forget. *arXiv:2305.17493*. https://arxiv.org/abs/2305.17493

Soergel, D. A. W. (2014). Rampant software errors may undermine scientific results. *F1000Research*, 3, 303. https://doi.org/10.12688/f1000research.5930.2

Stripelis, D., Gupta, U., Saleem, H., Dhinagar, N., Ghai, T., Anastasiou, R. C., Asghar, A., Ver Steeg, G., Ravi, S., Naveed, M., Thompson, P. M., & Ambite, J. L. (2022). Secure & private federated neuroimaging. *arXiv:2205.05249*. https://arxiv.org/abs/2205.05249

Taylor, P. A., Glen, D. R., Reynolds, R. C., Basavaraj, A., Moraczewski, D., & Etzel, J. (2023). Editorial: Demonstrating quality control (QC) procedures in fmri. *Frontiers in Neuroinformatics*, 17, 1205928. https://doi.org/10.3389/fnins.2023.1205928

The Turing Way Community. (2023). *The Turing Way: A handbook for reproducible, ethical and collaborative research* (Version 1.0.2). Zenodo. https://doi.org/10.5281/zenodo.7625728

Thomas, A., Ré, C., & Poldrack, R. (2022). Self-supervised learning of brain dynamics from broad neuroimaging data. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, & A. Oh (Eds.), *Advances in neural information processing systems* (Vol. 35, pp. 21255–21269). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2022/file/8600a9df1a087a9a66900cc8c948c3f0-Paper-Conference.pdf

Traut, N., Heuer, K., Lemaıtre, G., Beggiato, A., Germanaud, D., Elmaleh, M., Bethegnies, A., Bonnasse-Gahot, L., Cai, W., Chambon, S., Cliquet, F., Ghriss, A., Guigui, N., de Pierrefeu, A., Wang, M., Zantedeschi, V., Boucaud, A., van den Bossche, J., Kegl, B., … Varoquaux, G. (2022). Insights from an autism imaging biomarker challenge: Promises and threats to biomarker discovery. *NeuroImage*, 255, 119171. https://doi.org/10.1016/j.neuroimage.2022.119171

Tukey, J. W. (1962). The future of data analysis. *The Annals of Mathematical Statistics*, *33*(1), 1–67. https://doi.org/10.1214/aoms/1177704711

Tustison, N. J., Cook, P. A., Holbrook, A. J., Johnson, H. J., Muschelli, J., Devenyi, G. A., Duda, J. T., Das, S. R., Cullen, N. C., Gillen, D. L., Yassa, M. A., Stone, J. R., Gee, J. C., & Avants, B. B. (2021). The ANTsX ecosystem for quantitative biological and medical imaging. *Scientific Reports*, *11*(1), 9068. https://doi.org/10.1038/s41598-021-87564-6

van Viegen, T., Akrami, A., Bonnen, K., DeWitt, E., Hyafil, A., Ledmyr, H., Lindsay, G. W., Mineault, P., Murray, J. D., Pitkow, X., Puce, A., Sedigh-Sarvestani, M., Stringer, C., Achakulvisut, T., Alikarami, E., Atay, M. S., Batty, E., Erlich, J. C., Galbraith, B. V., … Peters, M. A. K. (2021). Neuromatch academy: Teaching computational neuroscience with global accessibility. *Trends in Cognitive Science*, *25*(7), 535–538. https://doi.org/10.1016/j.tics.2021.03.018

Veselovsky, V., Ribeiro, M. H., & West, R. (2023). Artificial artificial artificial intelligence: Crowd workers widely use large language models for text production tasks. *arXiv:2306.07899*. https://arxiv.org/abs/2306.07899

Woolston, C. (2022). Why science needs more research software engineers. *Nature: Career Q & A*. https://doi.org/10.1038/d41586-022-01516-2

Zador, A., Escola, S., Richards, B., Ölveczky, B., Bengio, Y., Boahen, K., Botvinick, M., Chklovskii, D., Churchland, A., Clopath, C., DiCarlo, J., Ganguli, S., Hawkins, J., Körding, K., Koulakov, A., LeCun, Y., Lillicrap, T., Marblestone, A., Olshausen, B., … Tsao, D. (2023). Catalyzing next-generation artificial intelligence through NeuroAI. *Nature Communications*, *14*(1), 1597. https://doi.org/10.1038/s41467-023-37180-x