
Review Articles and Meta-Analyses

Democratizing open neuroimaging: Neurodesk's approach to open data accessibility and utilization

Thuy T Dao, B.Eng^{1a}, Michèle Masson-Trottier, Ph.D¹, Fernanda L Ribeiro, Ph.D^{1,2}, Joshua Scarsbrook, B.Eng¹, Ashley Stewart, Ph.D¹, Xincheng Ye, M.D¹, Edan Hamilton, B.Eng¹, Steffen Bollmann, Ph.D^{1,3}^b, Aswin Narayanan, B.Eng^{4,5c}

¹ School of Electrical Engineering and Computer Science, The University of Queensland, ² Department of Medicine, Justus-Liebig-Universität Gießen, ³ Queensland Digital Health Centre, The University of Queensland, ⁴ Australian Institute for Bioengineering and Nanotechnology, The University of Queensland, ⁵ Australian National Imaging Facility

Keywords: data infrastructure, open science, data privacy, data standards, open data platforms, data ecosystems

<https://doi.org/10.52294/001c.144107>

Aperture Neuro

Vol. 5, Issue SI 2, 2025

Neuroimaging research has increasingly embraced open science practices, including the use and sharing of publicly available datasets. However, researchers still face significant challenges in accessing and preparing such datasets, from navigating through different storage systems and tools to complying with data privacy regulations. Moreover, the varied dependencies used by these tools often hinder reproducibility and accessibility, which in turn can negatively impact the reliability of analyses. In this paper, we show how Neurodesk streamlines both the use of open data and the preparation of datasets for sharing publicly. By integrating data access and standardized processing, Neurodesk maximizes the utility of open data initiatives while promoting accessibility and equity in the scientific community. We demonstrate Neurodesk's capabilities through real-world use cases and outline ongoing developments toward secure, scanner-integrated data processing, further supporting privacy-preserving collaboration in neuroscience.

INTRODUCTION

Neuroimaging data are crucial for studying brain structure and function and their relationship to human behaviour,¹⁻³ but acquiring high-quality data is costly and demands specialized expertise. To address these challenges and enable the pooling of datasets for larger and more comprehensive studies, two strategies have become common practice: (1) making smaller datasets openly available to the public through data repositories, like OpenNeuro⁴ and Open Science Framework (OSF)⁵; and (2) sharing data through large-scale consortium-led initiatives, such as the Human Connectome Project,⁶ the UKBiobank (UKBB) imaging,⁷ and the International Consortium for Brain Mapping (ICBM).⁸ Although it takes a great amount of effort to prepare these datasets, they enable the application of statistical models to derive biomarkers of neuropathologies with higher accuracy, reliability, and generalizability⁹⁻¹³ as they help control for confounding variables such as scanner type and pre-processing methods.

Beyond improving statistical model accuracy and generalizability, large neuroimaging datasets have also been used to create average templates, such as the MNI152 atlas,¹⁴⁻¹⁶ which serve as a common reference in many processing tools. Despite efforts made to include diverse ethnicities, genders, and age groups, these large-scale datasets often lack balanced demographic representation,^{14,15} leading to suboptimal performance for underrepresented populations.¹⁷ This highlights the potential of pooling smaller, more diverse datasets to complement large consortia efforts, ultimately making neuroimaging research more inclusive and equitable.

Accordingly, addressing demographic imbalances in neuroimaging studies requires greater international collaboration, but data sharing across borders comes with regulatory challenges. Many open data initiatives aim to protect personal information while enabling public access and compliance with diverse regulatory frameworks. These efforts typically involve removing personal identifiers and obtaining adequate consent from participants. Specifically,

^a Corresponding Author:

Thuy Thanh Dao
thanhthuy.dao@uq.edu.au

^b Co-Senior Author

^c Co-Senior Author

Open Brain Consent¹⁸ provides sample consent forms, template data user agreements tailored to specific regulations — such as the Health Insurance Portability and Accountability Act of 1996 (HIPAA) in the USA or the European General Data Protection Regulation (GDPR) in the EU — and recommended de-identification tools. Besides federal regulation compliance, there have been attempts to establish an international cross-border policy to govern data more efficiently to encourage more geographically diverse datasets.¹⁹ However, reaching a global consensus on principles that address legal, ethical, infrastructural, and financial issues requires significant effort and time. In the absence of such an agreement, researchers and their institutions are responsible for navigating these complexities, often limiting the scope and inclusivity of a shared dataset.

Moreover, publishing open datasets remains challenging in practice. To support the growing demand for data sharing, various repositories have emerged, each catering to different requirements. General-purpose repositories like OSF⁵ and Zenodo²⁰ offer flexible storage solutions for public and private datasets without enforcing strict data formats. Meanwhile, neuroimaging-specific repositories — such as OpenNeuro,⁴ BrainLife,²¹ The Image and Data Archive,²² the Canadian Open Neuroscience Platform,²³ and Neurobagel²⁴ — incorporate confidentiality safeguards, including privacy policies and secure data governance techniques. These systems facilitate responsible data sharing, but each repository is designed with different priorities, such as security, flexibility, interoperability, or accessibility. As a result, researchers must navigate varying data preparation requirements and carefully evaluate which repository aligns with their research needs and complies with institutional or national regulations.

Finally, even when datasets are openly available and pooled for larger studies, ensuring reproducible results when using them remains a critical challenge. Differences in the processing steps can introduce unintended biases for downstream analyses.²⁵ Such variability can compromise comparability across datasets and potentially undermine the benefits of data aggregation.

Altogether, these challenges complicate data sharing, hinder collaboration across research groups and raise compliance concerns with regulatory frameworks. To overcome these issues, Neurodesk²⁵ was developed as an open-source community-driven platform that offers a containerized data analysis environment to facilitate reproducible analysis of neuroimaging data. Here, we discuss and demonstrate how Neurodesk addresses these issues. Through real-world data usage and sharing scenarios, we highlight its role in improving collaboration and reproducibility in neuroimaging research:

1. By supporting the entire open data lifecycle — from preprocessing to data wrangling to publishing — Neurodesk ensures interoperability with different open data repositories and standardized tools in compliance with widely adopted standards like the Brain Imaging Data Structure (BIDS).

2. By providing a flexible infrastructure that supports both centralized and decentralized collaboration models to enable more publicly available datasets.
3. By enabling researchers to run analysis pipelines on any computing environment, it ensures that researchers can access and analyze datasets in a reproducible manner.

THE NEURODESK PLATFORM

NEURODESK - A SCALABLE PLATFORM FOR DATA CONSUMPTION AND PRODUCTION

Neuroimaging workflows require integrating multiple tools that often lead to compatibility and resource constraints. Neurodesk²⁵ addresses this challenge by providing a modular, lightweight, and scalable platform. Built on software containers, Neurodesk enables on-demand access to a comprehensive suite of neuroimaging tools, facilitating seamless processing, sharing, and access to datasets in an efficient and flexible environment (Figure 1).

To help with data preparation for open data repositories, Neurodesk brings tools to aid standardization, a critical step for facilitating collaboration across institutions. Neurodesk offers tools like dcm2niix,²⁶ heudiconv,²⁷ bidscoin,²⁸ and sovaBIDS²⁹ to standardize data into the BIDS format³⁰ to avoid manual organization, which is both time-consuming and error-prone. Once standardized, datasets can be seamlessly processed using BIDS-compatible tools such as QSMxT,³¹ BIDSApps,³² MRIQC,³³ fMRIPrep,³⁴ ASLprep³⁵ and many more. These pipelines ensure researchers can process and analyze data consistently and adhere to best practices.

Beyond neuroimaging software, Neurodesk also provides a range of tools to interface with various data sources — such as scanners, cloud storage solutions, and open data repositories. This simplifies data retrieval, ensuring that researchers can access datasets with minimal technical overhead.

Finally, Neurodesk simplifies open data sharing by integrating DataLad,³⁶ which enables data transfer to and from public repositories, like OSF⁵ or OpenNeuro.⁴ Additionally, Neurodesk supports various data transfer and cloud storage solutions, including Globus, OneDrive, Nextcloud, and ownCloud, allowing for secure and scalable collaboration across research institutions.

DATA COLLABORATION MODELS TO ACCOMMODATE VARIED DATA PRIVACY POLICIES

Neurodesk addresses the challenge of balancing openness and responsibility by supporting two models for data collaboration: **centralized** and **decentralized** (Figure 2). These collaboration models promote the use of open data and facilitate the legal and ethical sharing of processed derivatives to open repositories for large-scale, inclusive studies.

In the **centralized** collaboration model (Figure 2a), a Neurodesk-managed cloud instance allows collaborators to

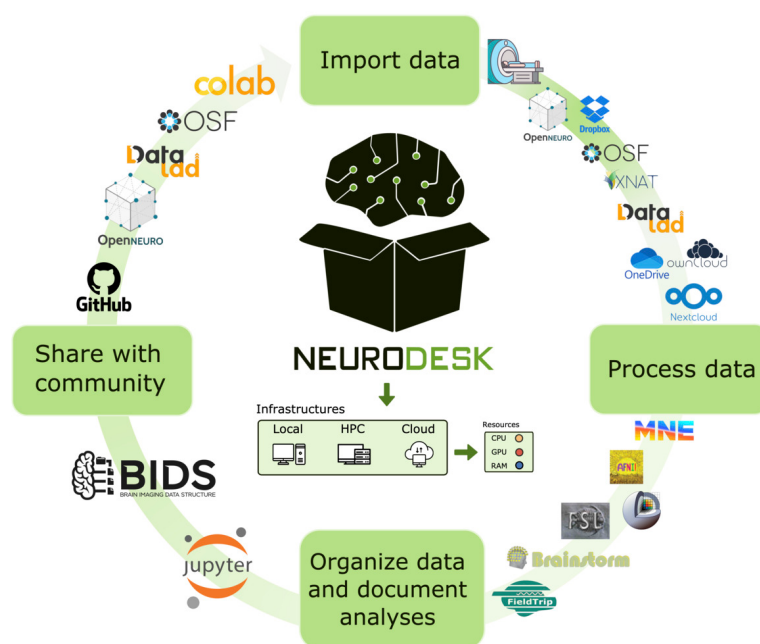


Figure 1. Data production and consumption workflows supported by Neurodesk with customizable resource allocation (CPU, GPU, RAM) and flexible deployment on local systems, HPC clusters, or cloud infrastructure.

access a shared storage environment, avoiding the need for each user to download and manage datasets individually. Data can be centrally retrieved from sources such as open data repositories or institutional storage systems and made available to all authorized users. Customizable disk space can be attached to the instance, allowing teams to allocate resources based on project needs. By centralizing data and tools, Neurodesk streamlines workflows while enforcing access control through authentication. A shared instance reduces entry barriers for users with limited computational resources or technical expertise and supports multi-site projects by synchronizing workflows and data analysis.

In the **decentralized** model (Figure 2b), separate Neurodesk installations are maintained at each institution, allowing researchers to process sensitive local data without sharing it with collaborators. This approach complies with stricter privacy policies where their data governance does not allow data sharing outside of the institution. Neurodesk's decentralized deployment allows institutions to maintain local control over sensitive datasets. Researchers can perform analyses on-premises, ensuring that sensitive data never leaves their institution's secure environment. Appropriately de-identified data can then be shared among collaborators and aggregated. It enables the creation of geographically diverse datasets while ensuring ethical and regulatory compliance for robust and generalizable research.

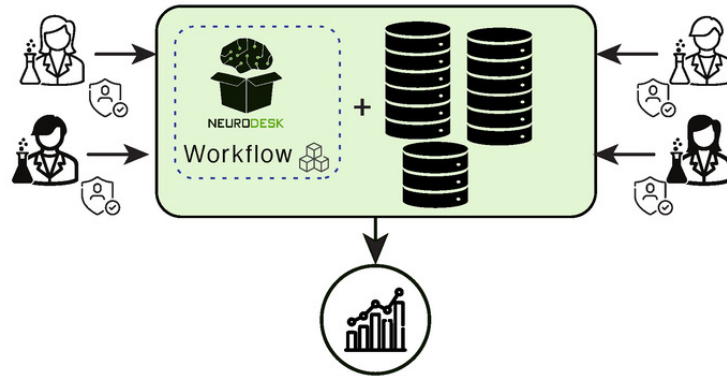
NEURODESK'S INFRASTRUCTURE FACILITATES DATA AND TOOL ACCESS FROM ANY COMPUTING ENVIRONMENT AND REPRODUCIBLE ANALYSIS

Neurodesk provides a containerized and modular computing environment that enables seamless access to neuroimaging tools and datasets across different computing platforms. Leveraging Docker³⁷/Podman³⁸ and Singularity³⁹/Apptainer⁴⁰, it offers an all-in-one solution where each tool is packaged with its dependencies to ensure portability and reproducibility. This setup eliminates compatibility issues and ensures consistent, reproducible analysis of open datasets across personal computers, high-performance computing (HPC) clusters, and cloud systems. This versatility supports both small-scale and large-scale computation, enabling research teams to collaborate across different infrastructures.

PRACTICAL APPLICATIONS OF NEURODESK: FROM OPEN DATA TO SCALABLE, PRIVACY-AWARE WORKFLOWS

To demonstrate how Neurodesk addresses key challenges in using open neuroimaging data — complex data processing workflows, cross-institutional collaboration, reproducibility, and accessibility — we present several example use cases demonstrating its impact on promoting the use and sharing of open datasets. These examples illustrate how researchers leverage Neurodesk to streamline complex data processing pipelines, collaborate securely across institutions, and scale reproducible analyses from local machines to HPC environments. Through these use cases, we high-

(a) Centralized collaboration



(b) Decentralized collaboration

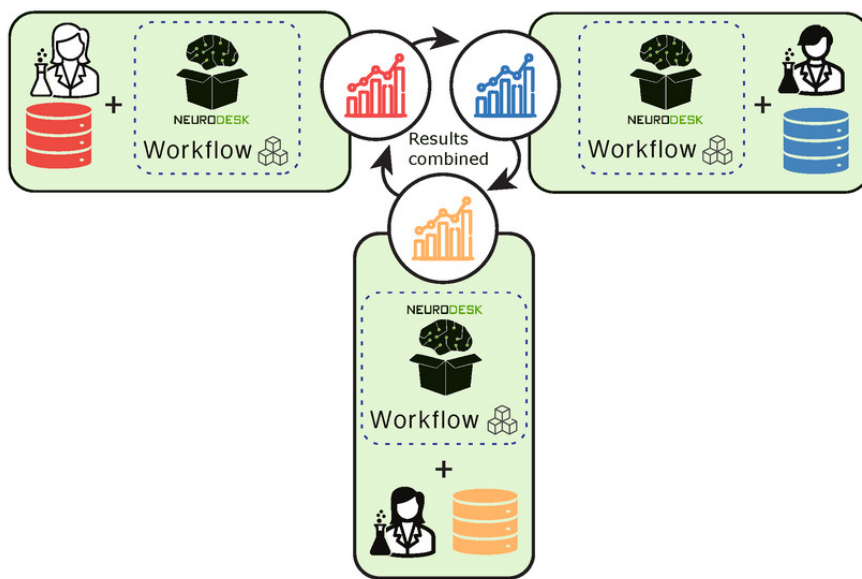


Figure 2. The (a) centralized and (b) decentralized collaboration models supported by Neurodesk offer flexibility to users for complying with different data-sharing policies.

light the platform's ability to facilitate the use and sharing of open datasets, enabling transparent and collaborative neuroimaging research.

USE CASE 1: PROCESSING BEHAVIOURAL AND NEUROIMAGING DATA

The first use case covers a study investigating relationships between brain structure, function, and behaviour, in which researchers collect behavioural and MRI data. In such studies, typically with smaller sample sizes, researchers handle all steps — from data acquisition at the MRI scanner to data storing, preprocessing, and analysis (Figure 3).

Neurodesk simplifies this process, providing a unified platform where essential neuroimaging tools are readily accessible. Researchers can perform preprocessing, visualization, and BIDS conversion without the burden of installing and configuring multiple tools. During data acquisition, researchers often perform quality checks before recruiting more participants. This involves converting raw

DICOM data from the MRI scanner into other formats for visualization. Neurodesk streamlines this process by supporting data conversion tools, like *dcm2nii*,²⁶ and visualization software, such as *MRICroGL*.⁴² As datasets grow, standardizing data structures becomes increasingly important for efficient processing and sharing. Neurodesk supports this through tools that facilitate BIDS formatting and quality assurance. Additionally, Python-based statistical and correlational analyses can be easily performed within the same environment, enhancing workflow continuity. To support open data sharing, Neurodesk also facilitates data de-identification, such as defacing brain images, prior to sharing datasets with public repositories. By integrating tools across the entire workflow, Neurodesk enables researchers to prepare, analyze, and share open neuroimaging data in a streamlined and reproducible manner.

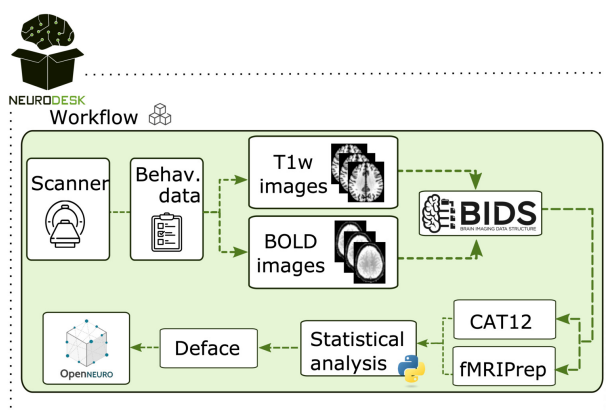


Figure 3. Use case 1 - In this scenario, the researcher uses Neurodesk to optimise the research process following data collection. After acquiring behavioural and imaging data, the imaging data are organized following BIDS specifications using either *BIDScoin*²⁸, *dcm2niix*, or *heudiconv*²⁷. Each tool has unique features, and depending on their needs, the researcher can choose the best-suited tool: *dcm2bids* can convert almost any set of DICOMs into NIFTIs with JSON sidecar files but requires additional steps for arranging the data in a BIDS structure; *heudiconv* is highly flexible but requires Python scripting to run a conversion; lastly, *BIDScoin* offers an intuitive and interactive GUI. Once the data is organized in a standardized manner, structural MRI data can be processed using the *CAT12*⁴¹ toolbox for voxel-based morphometry (VBM), while functional MRI data preprocessing can be performed with *fMRIprep*. Statistical analyses of processed data can then be conducted using Python-based tools. Finally, the researcher can decide to share the study data on OpenNeuro after defacing using *pydeface* or could also decide to store the data on any repository using *DataLad*³⁶.

USE CASE 2: OPEN-SCIENCE EDUCATION

The second use case is an educational workshop introducing neuroimaging tools. Workshops often require participants to access and analyze shared neuroimaging datasets, but setting up the required tools and downloading such datasets on individual devices can be time-consuming and impractical.

Neurodesk addresses this by enabling a shared cloud instance where all participants can securely access pre-configured neuroimaging tools and datasets in a standardized environment (Figure 4c). This ensures a smooth and consistent experience for running the workshop without delays. Every workshop participant is securely authenticated and works within an identical environment, making it very easy for the workshop demonstrators to guide learners through the material. The cloud-based environment also lowers technical barriers, allowing participants to join from devices that would otherwise be unsuitable for neuroimaging software, including tablets. Workshop demonstrators can access learner sessions via an admin panel, providing real-time guidance and troubleshooting without requiring

participants to configure software locally. Importantly, once the workshop is completed, the learners can install Neurodesk on their own computers or their institution's infrastructure and continue to use the tools they learned about during the workshop (Figure 4b). By centralizing the data and tools, Neurodesk is an ideal platform for education and training in the neuroimaging community.

USE CASE 3: PROCESSING OF SENSITIVE DATA FOR MULTI-SITE PROJECTS

The third use case addresses the challenge of privacy-compliant multi-institutional collaboration, where researchers from different institutes acquire data under similar scan protocols but need to analyze their datasets without sharing sensitive data due to HIPAA or GDPR restrictions. An additional challenge is to ensure that every collaborator runs exactly the same analysis with identical parameters so results can later be aggregated across sites, thereby increasing the generalizability of the methods.

To overcome this, Neurodesk enables a decentralized workflow, where researchers can run identical analyses on their machines while keeping sensitive datasets private (Figure 5). Instead of transferring identifiable or any kind of personal data — which may be legally restricted — collaborators can share only de-identified summary statistics and aggregate them in a meta-analysis. The ENIGMA consortium has exemplified that this approach allows for broader generalization of findings across different populations and sites without violating privacy regulations.

Without Neurodesk, ensuring consistency across institutions would be highly complex, as differences in software environments, dependencies, and computational setups could introduce variability in results.⁴³ This variability can undermine the statistical power of the analysis, leading to potentially unreliable conclusions. By providing a standardized and privacy-aware framework, Neurodesk removes these barriers, allowing institutions to collaborate securely while upholding strict data governance policies and ensuring the reproducibility and statistical integrity of their analyses.

USE CASE 4: SCALING ANALYSIS FROM PCS TO HIGH-PERFORMANCE COMPUTING CLUSTERS

The fourth use case highlights the challenge of developing a neuroimaging deep learning model that requires extensive preprocessing and access to large datasets, which often cannot be easily transferred due to data scale. In such experiments, researchers often begin by prototyping on local sample data before transitioning to running full-scale analyses on HPC or cloud-based resources where the complete dataset resides (Figure 6). Migrating analysis pipelines across environments introduces software compatibility issues and dependency conflicts, which hinder reproducibility.

Neurodesk provides a portable platform enabling flexible data access — either an open data repository or private storage — across multiple infrastructures. Researchers can begin their work using Neurodesk Play, a browser-based en-

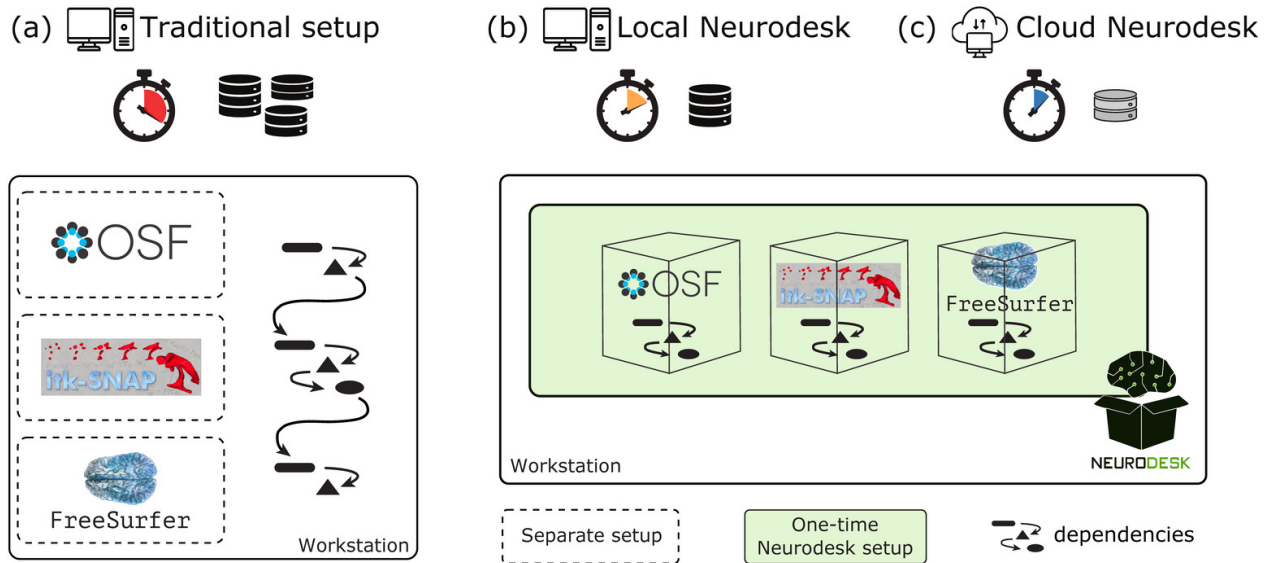


Figure 4. Use case 2 - This example illustrates different ways to set up a workshop focused on structural neuroimaging analysis using three essential tools: OSF, ITK-SNAP and FreeSurfer. **(a)** Students opting for a traditional setup face significant logistical hurdles; they must individually download each tool and its dependencies, consuming considerable time and storage space. Additionally, configuring and troubleshooting the workflows manually adds complexity and potential delays. In contrast, students utilizing the Neurodesk environment — whether accessed **(b)** locally or through **(c)** a centralized cloud instance — benefit from pre-configured, containerized tools, streamlining their setup experience. This containerized approach significantly reduces storage demands and eliminates dependency management challenges. Cloud-based access further accelerates this setup, providing instantaneous and uniform environments. Ultimately, Neurodesk enables a seamless, consistent, and efficient educational experience, fostering deeper engagement with neuroimaging tools.

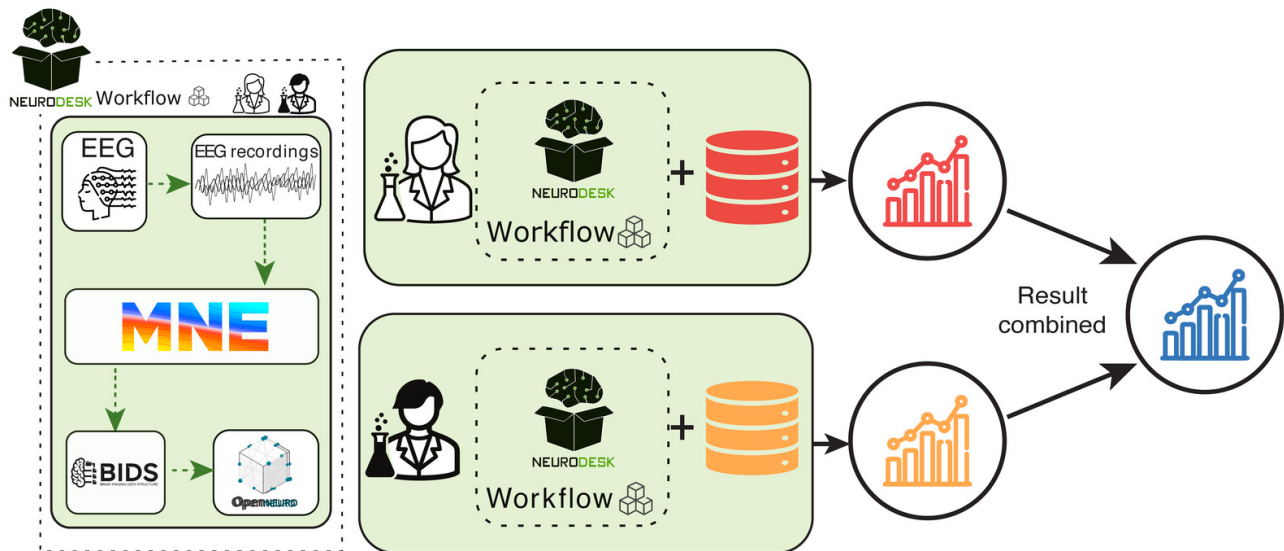


Figure 5. Use case 3 - This figure illustrates a collaboration where two researchers collect EEG data at their respective institutes. Due to differing privacy and data-sharing policies, they process the data locally using Neurodesk. The researchers develop a shared workflow, apply it to their local data, and then combine the results to generate the final project outcomes.

environment requiring no installation, and later transition to running the same analysis pipeline on a local machine, an HPC cluster, or a hosted Jupyter Notebook service. By enabling flexible data access and portable processing across

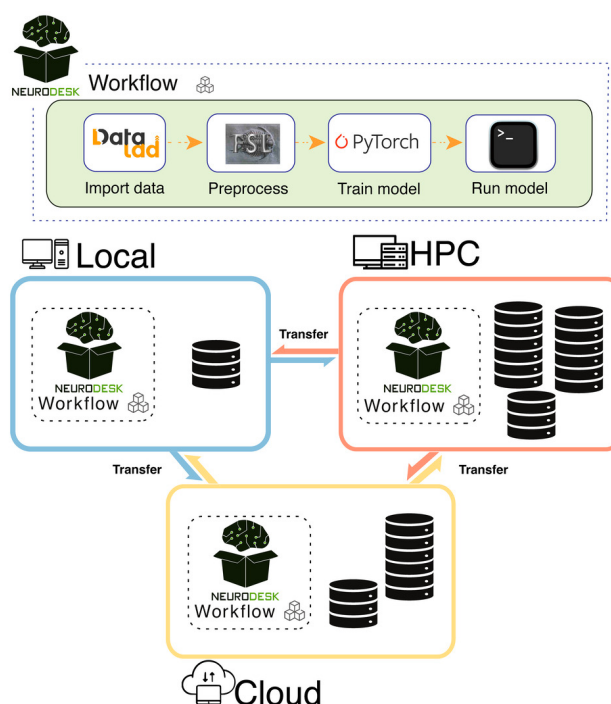


Figure 6. Use case 4 - This figure illustrates a workflow for developing a deep learning model for neuroimaging. The user loads large datasets, preprocesses images by registering and masking them using FSL, and then trains the model with PyTorch. Once trained, the user can share model checkpoints for others to run. The Neurodesk-based workflow, set up in a Jupyter Notebook, allows the user to shift between local machines and HPC for accessing different data subsets while having access to the tools for such a workflow without reinstalling tools.

infrastructures, Neurodesk facilitates collaboration and ensures that neuroimaging workflows remain scalable and efficient.

DISCUSSION

Here, we showcased how Neurodesk provides a portable, flexible, accessible platform that enhances open neuroimaging data sharing and usage. We provided guided examples on how researchers can work with and collaborate on different data sources, whether using open datasets or processing data locally. These use cases show that Neurodesk empowers researchers to prepare and publish standardized, shareable outputs to public or private repositories. We summarize the key advantages of using Neurodesk for neuroimaging experiments as follows:

- **Standardized data processing:** Neurodesk streamlines neuroimaging workflows by fully supporting BIDS-compliant data structuring and conversion tools, enabling researchers to organize datasets according to standardized community practices.
- **Secure and flexible data sharing:** Neurodesk provides both centralized and decentralized data-sharing models, ensuring compliance with international data protection regulations. By facilitating responsible data sharing, Neurodesk supports the development of more robust and generalizable statistical models, advancing neuroimaging research.

- **Portable analysis:** Neurodesk's modular design ensures a consistent environment across platforms for seamless dataset access and reproducible analysis. This makes collaborative projects, hackathons, and workshops more efficient by providing consistent execution environments to all participants.

Neurodesk has had a significant impact on workshops, training programs, and research projects worldwide, particularly for researchers with limited resources or those with limited permissions on institutional workstations. By providing a portable, ready-to-use platform, Neurodesk lowers technical barriers, enabling more equitable access to advanced neuroimaging tools. In the rapidly evolving neuroimaging research field, Neurodesk stands out by swiftly integrating new tools and adapting to emerging research needs. Its commitment to open science and technological innovation makes it a crucial resource for the global neuroimaging community.

Beyond facilitating collaboration, Neurodesk is actively advancing its development to address the challenges associated with data security and regulatory compliance in neuroimaging research. Recognizing the need for faster translation of research pipelines to clinical questions, Neurodesk is expanding its toolset to be accessible directly on imaging scanners. This is currently already possible on Siemens scanners through the Open Recon framework.⁴⁴ By leveraging Open Recon, a platform that allows image processing containers to run directly on the scanner, Neurodesk containers can be run on the scanner and fully integrate sci-

entific processing workflows into clinical workflows. This innovative approach allows researchers to execute the entire workflow — from data collection to processing — on the scanner itself, ensuring that sensitive data never leaves the secure environment. As Neurodesk continues to evolve, these efforts exemplify its commitment to creating a seamless platform for open collaboration, ultimately enabling clinical translation.

.....

DATA AND CODE AVAILABILITY

Open-source code for the development of Neurodesk infrastructure is available on GitHub (<https://github.com/neurodesk>).

FUNDING SOURCES

The authors acknowledge funding through an ARC Linkage grant (LP200301393). This research was supported by the

use of the ARDC Nectar Research Cloud, a collaborative Australian research platform supported by the NCRIS-funded Australian Research Data Commons (ARDC), the EGI Federation with support from CESNET-MCC and Indiana University's Jetstream2 Cloud. This work is supported by the Wellcome Trust with a Discretionary Award as part of the Chan Zuckerberg Initiative (CZI), The Kavli Foundation, and Wellcome's Essential Open Source Software for Science (Cycle 6) Program (Grant Ref: [313306/Z/24/Z]). FLR acknowledges support through the European Union's Horizon Europe research and innovation funding program under the Marie Skłodowska-Curie Actions project ID 101146996.

CONFLICTS OF INTEREST

The authors have no competing interests to declare that are relevant to the content of this article.

Submitted: April 28, 2025 CDT. Accepted: August 14, 2025 CDT. Published: September 26, 2025 CDT.



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license's legal deed at <http://creativecommons.org/licenses/by/4.0> and legal code at <http://creativecommons.org/licenses/by/4.0/legalcode> for more information.

REFERENCES

1. Poldrack RA, Gorgolewski KJ. Making big data open: Data sharing in neuroimaging. *Nature Neuroscience*. 2014;17(11):1510-1517. doi:[10.1038/nn.3818](https://doi.org/10.1038/nn.3818)
2. Poline JB, Breeze JL, Ghosh S, et al. Data sharing in neuroimaging research. *Frontiers in Neuroinformatics*. 2012;6. doi:[10.3389/fninf.2012.00009](https://doi.org/10.3389/fninf.2012.00009)
3. the Alzheimer's Disease Neuroimaging Initiative, EPIGEN Consortium, IMAGEN Consortium, Saguenay Youth Study (SYS) Group, Thompson PM, Stein JL, et al. The ENIGMA Consortium: Large-scale collaborative analyses of neuroimaging and genetic data. *Brain Imaging and Behavior*. 2014;8(2):153-182. doi:[10.1007/s11682-013-9269-5](https://doi.org/10.1007/s11682-013-9269-5)
4. Markiewicz CJ, Gorgolewski KJ, Feingold F, et al. The OpenNeuro resource for sharing of neuroscience data. *eLife*. 2021;10:e71774. doi:[10.7554/eLife.71774](https://doi.org/10.7554/eLife.71774)
5. Foster ED, Deardorff A. Open Science Framework (OSF). *Journal of the Medical Library Association*. 2017;105(2). doi:[10.5195/jmla.2017.88](https://doi.org/10.5195/jmla.2017.88)
6. Van Essen DC, Smith SM, Barch DM, Behrens TEJ, Yacoub E, Ugurbil K. The WU-Minn Human Connectome Project: An overview. *NeuroImage*. 2013;80:62-79. doi:[10.1016/j.neuroimage.2013.05.041](https://doi.org/10.1016/j.neuroimage.2013.05.041)
7. Conroy M, Sellors J, Effingham M, et al. The advantages of UK Biobank's open-access strategy for health research. *Journal of Internal Medicine*. 2019;286(4):389-397. doi:[10.1111/joim.12955](https://doi.org/10.1111/joim.12955). PMID:31283063
8. Mazziotta JC, Toga AW, Evans A, Fox P, Lancaster J. A Probabilistic Atlas of the Human Brain: Theory and Rationale for Its Development. *NeuroImage*. 1995;2(2):89-101. doi:[10.1006/nimg.1995.1012](https://doi.org/10.1006/nimg.1995.1012)
9. Horien C, Noble S, Greene AS, et al. A hitchhiker's guide to working with large, open-source neuroimaging datasets. *Nature Human Behaviour*. 2020;5(2):185-193. doi:[10.1038/s41562-020-01005-4](https://doi.org/10.1038/s41562-020-01005-4)
10. Smith SM, Nichols TE. Statistical Challenges in "Big Data" Human Neuroimaging. *Neuron*. 2018;97(2):263-268. doi:[10.1016/j.neuron.2017.12.018](https://doi.org/10.1016/j.neuron.2017.12.018)
11. Noble S, Scheinost D, Constable RT. Cluster failure or power failure? Evaluating sensitivity in cluster-level inference. *NeuroImage*. 2020;209:116468. doi:[10.1016/j.neuroimage.2019.116468](https://doi.org/10.1016/j.neuroimage.2019.116468)
12. Bzdok D, Yeo BTT. Inference in the age of big data: Future perspectives on neuroscience. *NeuroImage*. 2017;155:549-564. doi:[10.1016/j.neuroimage.2017.04.061](https://doi.org/10.1016/j.neuroimage.2017.04.061)
13. Schulz MA, Yeo BTT, Vogelstein JT, et al. Different scaling of linear models and deep learning in UKBiobank brain images versus machine-learning datasets. *Nature Communications*. 2020;11(1):4238. doi:[10.1038/s41467-020-18037-z](https://doi.org/10.1038/s41467-020-18037-z)
14. Fonov V, Evans AC, Botteron K, Almli CR, McKinstry RC, Collins DL. Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage*. 2011;54(1):313-327. doi:[10.1016/j.neuroimage.2010.07.033](https://doi.org/10.1016/j.neuroimage.2010.07.033)
15. Mazziotta J, Toga A, Evans A, et al. A four-dimensional probabilistic atlas of the human brain. *Journal of the American Medical Informatics Association: JAMIA*. 2001;8(5):401-430. doi:[10.1136/jamia.2001.0080401](https://doi.org/10.1136/jamia.2001.0080401). PMID:11522763
16. Fonov V, Evans A, McKinstry R, Almli C, Collins D. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*. 2009;47:S102. doi:[10.1016/S1053-8119\(09\)70884-5](https://doi.org/10.1016/S1053-8119(09)70884-5)
17. Gao P, Dong HM, Liu SM, et al. A Chinese multi-modal neuroimaging data release for increasing diversity of human brain mapping. *Scientific Data*. 2022;9(1):286. doi:[10.1038/s41597-022-01413-3](https://doi.org/10.1038/s41597-022-01413-3)
18. Bannier E, Barker G, Borghesani V, et al. The Open Brain Consent: Informing research participants and obtaining consent to share brain imaging data. *Human Brain Mapping*. 2021;42(7):1945-1951. doi:[10.1002/hbm.25351](https://doi.org/10.1002/hbm.25351)
19. Eke DO, Bernard A, Bjaalie JG, et al. International data governance for neuroscience. *Neuron*. 2022;110(4):600-612. doi:[10.1016/j.neuron.2021.11.017](https://doi.org/10.1016/j.neuron.2021.11.017). PMID:34914921
20. European Organization For Nuclear Research, OpenAIRE. Zenodo. Published online 2013. doi:[10.25495/7GXXK-RD71](https://doi.org/10.25495/7GXXK-RD71)
21. Hayashi S, Caron BA, Heinsfeld AS, et al. Brainlife.io: A decentralized and open-source cloud platform to support neuroscience research. *Nature Methods*. 2024;21(5):809-813. doi:[10.1038/s41592-024-02237-2](https://doi.org/10.1038/s41592-024-02237-2)

22. Neu SC, Crawford KL, Toga AW. The image and data archive at the laboratory of neuro imaging. *Frontiers in Neuroinformatics*. 2023;17:1173623. doi:[10.3389/fninf.2023.1173623](https://doi.org/10.3389/fninf.2023.1173623)
23. Harding RJ, Bermudez P, Bernier A, et al. The Canadian Open Neuroscience Platform—An open science framework for the neuroscience community. Papin JA, ed. *PLOS Computational Biology*. 2023;19(7):e1011230. doi:[10.1371/journal.pcbi.1011230](https://doi.org/10.1371/journal.pcbi.1011230)
24. Urchs S, Dai A, Armoza J, Jahanpour A, Bhagwat N, Poline JB. Neurobagel CLI: Command line tool for Neurobagel data parsing and annotation. Published online June 2023. doi:[10.5281/ZENODO.8088210](https://doi.org/10.5281/ZENODO.8088210)
25. Renton AI, Dao TT, Johnstone T, et al. Neurodesk: An accessible, flexible and portable data analysis environment for reproducible neuroimaging. *Nature Methods*. Published online January 2024. doi:[10.1038/s41592-023-02145-x](https://doi.org/10.1038/s41592-023-02145-x)
26. Li X, Morgan PS, Ashburner J, Smith J, Rorden C. The first step for neuroimaging data analysis: DICOM to NIfTI conversion. *Journal of Neuroscience Methods*. 2016;264:47-56. doi:[10.1016/j.jneumeth.2016.03.001](https://doi.org/10.1016/j.jneumeth.2016.03.001)
27. Halchenko YO, Goncalves M, Ghosh S, et al. HeuDiConv — flexible DICOM conversion into structured directory layouts. *Journal of Open Source Software*. 2024;9(99):5839. doi:[10.21105/joss.05839](https://doi.org/10.21105/joss.05839)
28. Zwiers MP, Moia S, Oostenveld R. BIDScoin: A User-Friendly Application to Convert Source Data to Brain Imaging Data Structure. *Frontiers in Neuroinformatics*. 2022;15:770608. doi:[10.3389/fninf.2021.770608](https://doi.org/10.3389/fninf.2021.770608)
29. Mantilla Ramos YJ, Hoyos Madera BA, Bollmann S, et al. Sovabids: EEG-to-BIDS conversion software focused on automation, reproducibility and interoperability. In: ; 2022. <https://github.com/yimantilla/sofabids>
30. Gorgolewski KJ, Auer T, Calhoun VD, et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*. 2016;3(1):160044. doi:[10.1038/sdata.2016.44](https://doi.org/10.1038/sdata.2016.44)
31. Stewart AW, Robinson SD, O'Brien K, et al. QSMxT: Robust masking and artifact reduction for quantitative susceptibility mapping. *Magnetic Resonance in Medicine*. 2022;87(3):1289-1300. doi:[10.1002/mrm.29048](https://doi.org/10.1002/mrm.29048)
32. Gorgolewski KJ, Alfaro-Almagro F, Auer T, et al. BIDS apps: Improving ease of use, accessibility, and reproducibility of neuroimaging data analysis methods. Schneidman D, ed. *PLOS Computational Biology*. 2017;13(3):e1005209. doi:[10.1371/journal.pcbi.1005209](https://doi.org/10.1371/journal.pcbi.1005209)
33. Esteban O, Birman D, Schaer M, Koyejo OO, Poldrack RA, Gorgolewski KJ. MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites. Bernhardt BC, ed. *PLOS ONE*. 2017;12(9):e0184661. doi:[10.1371/journal.pone.0184661](https://doi.org/10.1371/journal.pone.0184661)
34. Esteban O, Markiewicz CJ, Blair RW, et al. fMRIPrep: A robust preprocessing pipeline for functional MRI. *Nature Methods*. 2019;16(1):111-116. doi:[10.1038/s41592-018-0235-4](https://doi.org/10.1038/s41592-018-0235-4)
35. Adebimpe A, Bertolero M, Dolui S, et al. ASLPrep: A platform for processing of arterial spin labeled MRI and quantification of regional brain perfusion. *Nature Methods*. 2022;19(6):683-686. doi:[10.1038/s41592-022-01458-7](https://doi.org/10.1038/s41592-022-01458-7)
36. Halchenko YO, Meyer K, Poldrack B, et al. DataLad: Distributed system for joint management of code, data, and their relationship. *Journal of Open Source Software*. 2021;6(63):3262. doi:[10.21105/joss.03262](https://doi.org/10.21105/joss.03262)
37. Merkel D. Docker: Lightweight Linux containers for consistent development and deployment. *Linux J*. 2014;2014(239).
38. Heon M et. al, Walsh D, Baude B, et al. Podman - : A tool for managing OCI containers and pods. Published online January 2018. doi:[10.5281/ZENODO.4735634](https://doi.org/10.5281/ZENODO.4735634)
39. Kurtzer GM, Sochat V, Bauer MW. Singularity: Scientific containers for mobility of compute. Gursoy A, ed. *PLOS ONE*. 2017;12(5):e0177459. doi:[10.1371/journal.pone.0177459](https://doi.org/10.1371/journal.pone.0177459)
40. Dykstra D. Apptainer Without Setuid. Published online September 2023. doi:[10.48550/arXiv.2208.12106](https://doi.org/10.48550/arXiv.2208.12106)
41. Gaser C, Dahnke R, Thompson PM, Kurth F, Luders E, the Alzheimer's Disease Neuroimaging Initiative. CAT: A computational anatomy toolbox for the analysis of structural MRI data. *GigaScience*. 2024;13:giae049. doi:[10.1093/gigascience/giae049](https://doi.org/10.1093/gigascience/giae049)
42. Rorden C. From MRICro to MRICron: The evolution of neuroimaging visualization tools. *Neuropsychologia*. 2025;207:109067. doi:[10.1016/j.neuropsychologia.2025.109067](https://doi.org/10.1016/j.neuropsychologia.2025.109067)

43. Glatard T, Lewis LB, Ferreira da Silva R, et al. Reproducibility of neuroimaging analyses across operating systems. *Frontiers in Neuroinformatics*. 2015;9. doi:[10.3389/fninf.2015.00012](https://doi.org/10.3389/fninf.2015.00012)

44. Wallace TE, Morales MA, Assana S, Demirel OB, Chow K, Nezafat R. Accelerating Cardiac MRI with Inline Generative AI and Open Recon. 2024;MAGNETOM Flash(87).