bids2ebrains: Automating BIDS Dataset Registration into the EBRAINS Knowledge Graph

Renqing Cuomao[1*], Raphaël Gazzotti[2*], Sophia Pieschnik[3], Oliver Schmid[4], Peyman Najafi[5], Andrew P. Davison[5], Cyril Pernet[6], Lyuba Zehl[2*]

[1]EPFL, Lausanne, Switzerland

[2]EBRAINS AISBL, Brussels, Belgium

[3]Institute of Basic Medical Sciences, Faculty of Medicine, University of Oslo, Oslo, Norway

[4]CSCS, ETH Zurich, Zurich, Switzerland

[5]Institut des Neurosciences Paris-Saclay, Université Paris-Saclay, CNRS, Saclay, France

[6]Neurobiology Research Unit, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark

* renqing.cumao@epfl.ch; raphael.gazzotti@ebrains.eu; lyuba.zehl@ebrains.eu

INTRODUCTION/MOTIVATION

The Brain Imaging Data Structure (BIDS) standard [1] and the openMINDS metadata framework [2] are cornerstones for findable, accessible, interoperable, and reusable (FAIR) neuroscience [3]. BIDS ensures consistency in dataset organization, while openMINDS enables explicit, ontology-driven metadata representations within the EBRAINS Knowledge Graph (KG) [4]. However, registering BIDS datasets in EBRAINS remains a mostly manual and error-prone process that requires detailed knowledge of metadata schemas. This hinders the large-scale integration and reuse of neuroscience datasets within EBRAINS. To overcome this obstacle, we first developed the bids2openminds converter [5] which is now complemented with the bids2ebrains [6].

METHODS

bids2ebrains facilitates the registration of BIDS datasets into the EBRAINS KG by providing a semi-automated pipeline for uploading openMINDS compliant metadata formatted as JSON-LD. The tool is available as (1) a command-line interface (CLI), (2) an interactive Streamlit-based graphical user interface (UI), and (3) a Python library, accommodating both technical and non-technical users. All interfaces enable the following workflow:

- **Conversion module** – relies on the existing prototype of the bids2openminds converter for translating BIDS metadata to openMINDS compliant JSON-LD files.
- **Scan module** – automatically detects and reports missing metadata entries expected by the EBRAINS curation service.
- **Patch module** – enables completion of missing fields either automatically (e.g., for file sizes) or interactively (e.g., for missing controlled vocabulary).

- Upload module – authenticates provider for uploading openMINDS compliant metadata of the BIDS dataset into an eligible EBRAINS KG space.

Once uploaded to an eligible EBRAINS KG space, metadata only requires final validation by the EBRAINS curation service before being released.

RESULTS AND DISCUSSION

We validated bids2ebrains using the ds001 open dataset from the BIDS examples repository [7]. Running the conversion module generated the essential set of openMINDS JSON-LD files for the Dataset, DatasetVersion, Subject, and File entities. The scan module identified all missing metadata entries across the detected files and automatically aggregated these entries into prompts. The interface of the patch module enabled the interactive completion of missing controlled vocabularies and free-text fields. Moreover it auto-completes fields that are inferable from the source data. The finally complete set of openMINDS metadata was then pushed into a dedicated EBRAINS KG test space via the upload module, after a dry-run for ensuring safe validation before final submission. Initial testing showed that the tool substantially reduces the effort needed for registration and curation when submitting BIDS compliant datasets to EBRAINS. These findings highlight how bids2ebrains streamlines data sharing and provide the basis for further discussion of its broader impact.

bids2ebrains enables the seamless integration of BIDS datasets into the EBRAINS KG, bridging community standards (BIDS) with infrastructure-level metadata frameworks (openMINDS). By combining automation, validation, and user-friendly interfaces, the tool lowers technical barriers and the required effort for dataset registration and curation. Looking ahead, bids2ebrains will enable both individual BIDS datasets registration and large-scale synchronization with external BIDS-compliant repositories (e.g., publicnEUro), provided that EBRAINS registration requirements are fulfilled, as expected/ensured by the tool's validation workflow. Its modular design enables future extensions to new modalities and metadata classes as BIDS and openMINDS evolve, positioning it as a key infrastructure tool for open and FAIR neuroscience, standardization alignment, and reproducible (meta)data registration and curation.

Keywords: BIDS, openMINDS, FAIR data, EBRAINS, Knowledge Graph, metadata curation, automation, interoperability, neuroinformatics, data integration

ACKNOWLEDGEMENTS

This work was developed as part of the Google Summer of Code 2025 program, under the mentorship of the International Neuroinformatics Coordinating Facility (INCF). This work also received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101147319 (EBRAINS 2.0).

REFERENCES

[1] Gorgolewski K, Auer T, Calhoun V, et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Sci Data. 2016;3:160044. https://doi.org/10.1038/sdata.2016.44

[2] Open Metadata Initiative. openMINDS metadata framework. Resource Identification Portal: RRID:SCR_023173. FAIRsharing.org: https://doi.org/10.25504/FAIRsharing.6ac6aa. See also GitHub repository (published 2023; updated 2025; accessed 2025-08-25): https://github.com/openMetadataInitiative/openMINDS

[3] Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016;3:160018. https://doi.org/10.1038/sdata.2016.18

[4] EBRAINS. EBRAINS Knowledge Graph. Resource Identification Portal. RRID:SCR_017612. See also (accessed 2025-08-26): https://kg.ebrains.eu

[5] Open Metadata Initiative. bids2openminds. (published 2024; updated 2025; accessed 2025-08-26) https://github.com/openMetadataInitiative/bids2openminds

[6] EBRAINS. bids2ebrains. (published 2025; accessed 2025-08-26) https://gitlab.ebrains.eu/ri/projects-and-initiatives/bids2ebrains/bids2ebrains

[7] Brain Imaging Data Structure. bids-examples. (published 2016; updated 2025; accessed 2025-08-26) https://github.com/bids-standard/bids-examples