

DEVELOPING A SYSTEM FOR THE EXTRACTION OF MULTIMODAL NEUROIMAGING SUB-DATASETS FROM STRUCTURED OR UNSTRUCTURED DATASETS

Sabrina Castro Nakamura¹; Yossi Zana^{2*}

1. Sabrina Castro Nakamura. Centro de Matemática, Computação e Cognição. Universidade Federal do ABC, São Bernardo do Campo, São Paulo, Brasil. sabrina.nakamura@aluno.ufabc.edu.br (ORCID: 0009-0005-2652-806X)
 2. Yossi Zana. Centro de Matemática, Computação e Cognição. Universidade Federal do ABC, São Bernardo do Campo, São Paulo, Brasil. yossi.zana@ufabc.edu.br (ORCID: 0000-0003-0539-7492)
- Autor apresentador

Palavras chaves: *EEG*; *fNIRS*; *TMS*; desenvolvimento de software; Neuroinformática.

Introdução: A reutilização de dados de neuroimagem, como a eletroencefalografia (EEG) e a espectroscopia funcional no infravermelho próximo (fNIRS), é um desafio devido à falta de padronização estrutural e de metadados nos repositórios atuais. Nesse contexto, foi desenvolvido um sistema automatizado para identificar, mapear e extrair subconjuntos de dados clinicamente relevantes.

Objetivos: Desenvolver e validar um software para a extração automatizada de subconjuntos de dados neurofisiológicos multimodais. O sistema deve unificar metadados heterogêneos, aplicar filtros de extração estruturais e exportar os resultados em formatos interoperáveis, estritamente alinhados ao padrão BIDS. A ferramenta deve apresentar uma interface gráfica intuitiva para a seleção de participantes, trials e canais, além de realizar a validação automática da integridade dos dados, a conversão entre diferentes formatos de EEG (BrainVision, EEGLAB, GDF, EDF, FIFF) e a geração de relatórios detalhados. Isso facilita análises científicas com rigor metodológico e conformidade normativa.

Materiais e métodos: O estudo utilizou uma base de dados pública (<https://github.com/OpenNeuroDatasets/ds003810>), dispensando a aprovação de um Comitê de Ética. O protótipo foi implementado em Python, apresentando uma interface gráfica modular em PyQt6 e processamento em segundo plano (backend). Integrou-se a biblioteca MNE-Python para a conversão unificada de cinco formatos neurofisiológicos e o PyBIDS para a validação hierárquica. Os testes funcionais avaliaram o carregamento e a filtragem estrutural, o controle de qualidade via inspeção visual com correção de linha de base e a configuração de exportação. Verificou-se uma extração segura, que preserva os arquivos originais e gera metadados automaticamente, garantindo a conformidade com o padrão BIDS.

Resultados/discussão: O sistema extraiu os dados neurofisiológicos com eficácia, identificando automaticamente a estrutura BIDS na interface principal (Fig. 1). A inspeção visual permitiu o isolamento de traçados ruidosos (Fig. 2), enquanto o sistema alertou ativamente sobre os riscos de conversão nos cinco formatos suportados. As subamostras foram exportadas com 100% de preservação dos dados brutos e geração automática de diretórios (Fig. 3). A estabilidade do software foi garantida pelo processamento em segundo plano e pelos registros de validação. Avaliada por métricas computacionais e conformidade com o BIDS, a ferramenta automatiza o fluxo laboratorial.

Conclusão: O sistema proposto cumpriu o objetivo de automatizar a manipulação de dados neurofisiológicos. A ferramenta demonstrou eficácia ao assegurar a preservação dos sinais originais e a conformidade estrutural com o padrão BIDS. Conclui-se que essa tecnologia otimiza o fluxo de trabalho laboratorial, mitiga falhas humanas na manipulação de arquivos e estabelece a base de dados estruturada necessária para a futura aplicação de inteligência artificial na neurociência.

Financiamento: Este projeto foi realizado com o apoio financeiro da Universidade Federal do ABC (UFABC), por meio de bolsa concedida pelo Programa de Desenvolvimento de Pesquisa Dirigida - Iniciação Tecnológica e Industrial (PDPD-ITI), vinculado ao Edital ProPes nº 19/2025 (Protocolo nº 23006.026827/2025-55).

Conflito de interesse: Não há existência de qualquer conflito de interesse, de natureza financeira relacionada à concepção, condução ou publicação deste trabalho.

Referências:

- 1.Appelhof S, et al. MNE-BIDS: Organizing electrophysiological data into the BIDS format and facilitating their analysis. J Open Source Softw. 2019;4(44):1896. Available from: <https://doi.org/10.21105/joss.01896>
- 2.Gorgolewski KJ, et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Sci Data. 2016;3:160044. Available from: <https://doi.org/10.1038/sdata.2016.44>
- 3.Luke R, et al. fNIRS-BIDS, the brain imaging data structure extension for functional near-infrared spectroscopy. Sci Data. 2021;8:299. Available from: <https://doi.org/10.1038/s41597-021-01022-2>
- 4.Nichols TE, et al. Best practices in data analysis and sharing in neuroimaging using MRI. Nat Neurosci. 2017;20:299-303. Available from: <https://doi.org/10.1038/nn.4500>
- 5.Poldrack RA, et al. Scanning the horizon: towards transparent and reproducible neuroimaging research. Nat Rev Neurosci. 2017;18:115-126. Available from: <https://doi.org/10.1038/nrn.2016.167>
- 6.Pernet CR, et al. EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. Sci Data. 2019;6:103. Available from: <https://doi.org/10.1038/s41597-019-0104-8>
- 7.Wilkinson MD, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016;3:160018. Available from: <https://doi.org/10.1038/sdata.2016.18>
- 8.Yucel MA, et al. Best practices for fNIRS publications. Neurophotonics. 2021;8(1):012101. Available from: <https://doi.org/10.1117/1.NPh.8.1.012101>

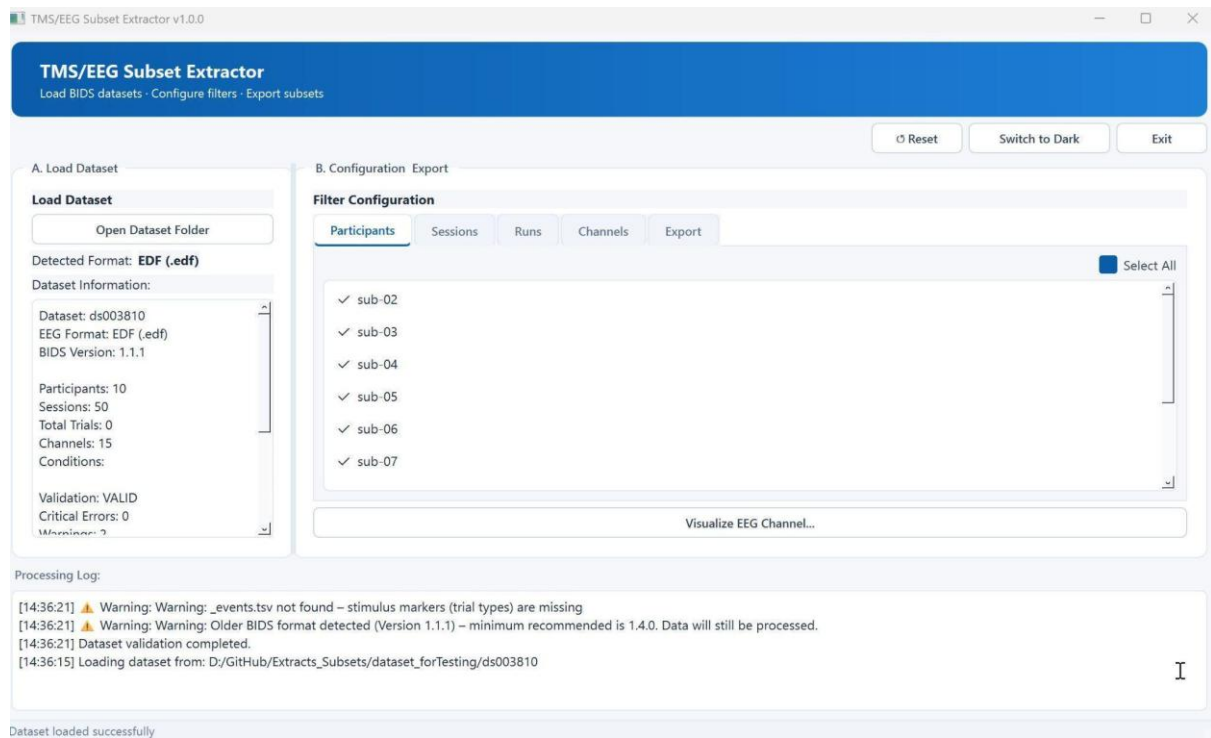


Figura 1: Interface gráfica principal do usuário, exibindo a etapa de carregamento, o registro de processamento com avisos de validação estrutural e a aba de filtragem de participantes.

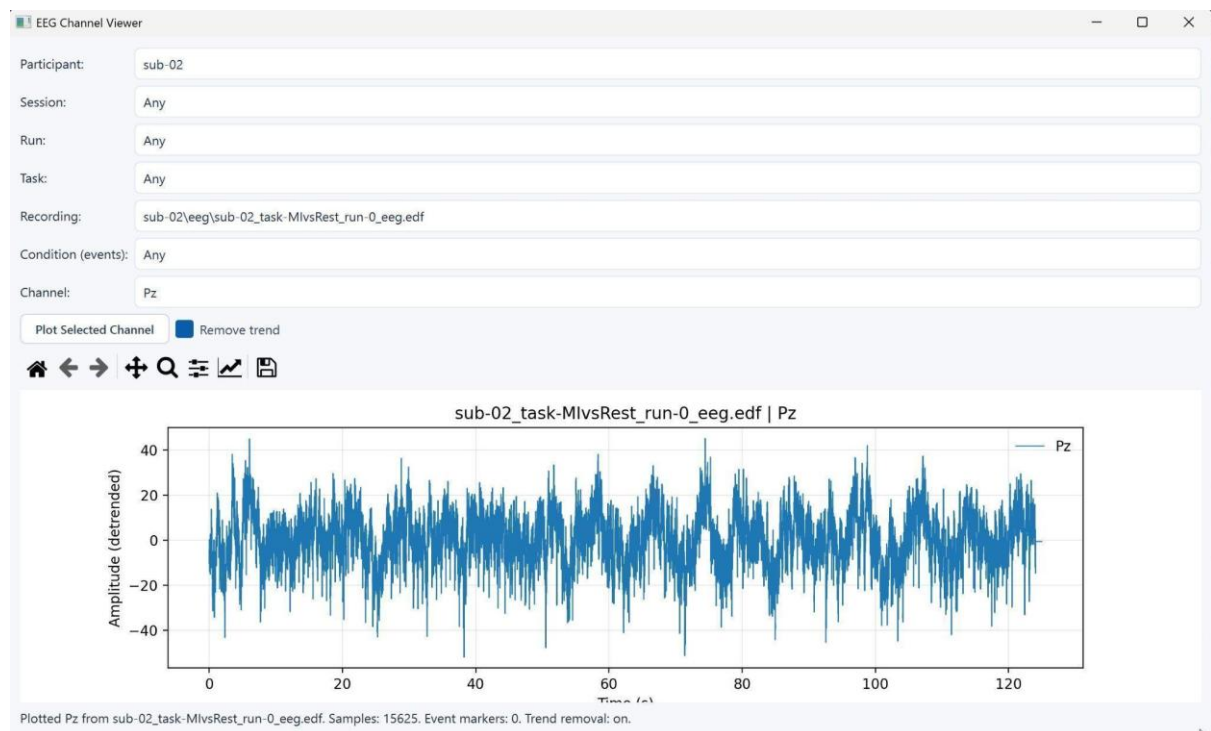


Figura 2: O visualizador de canais de EEG integrado realiza o controle de qualidade plotando um sinal bruto sem tendência para um participante e canal específicos.

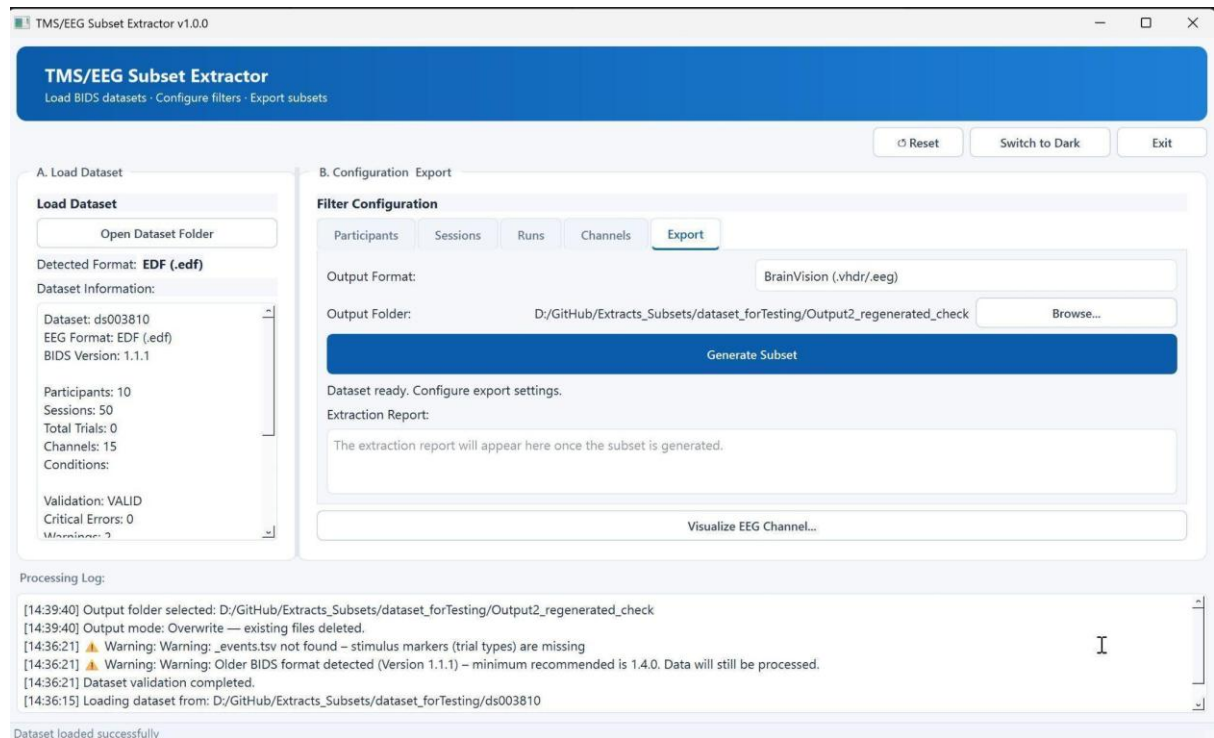


Figura 3: A aba de configuração de exportação permite ao usuário selecionar o formato de saída final e a pasta de destino para gerar o subconjunto personalizado.