NeuroImage 245 (2021) 118713

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

WeBrain: A web-based brainformatics platform of computational ecosystem for EEG big data analysis

Li Dong[a,b,d], Jianfu Li[a,b,d], Qiunan Zou[a], Yufan Zhang[a], Lingling Zhao[a], Xin Wen[a], Jinnan Gong[a], Fali Li[a], Tiejun Liu[a,b,d], Alan C. Evans[e], Pedro A. Valdes-Sosa[a,f], Dezhong Yao[a,b,c,d,*]

[a] The Clinical Hospital of Chengdu Brain Science Institute, MOE Key Lab for Neuroinformation, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu, China
[b] Research Unit of NeuroInformation, Chinese Academy of Medical Sciences, Chengdu 2019RU035, China
[c] School of Electrical Engineering, Zhengzhou University, Zhengzhou 450001, China
[d] Sichuan Institute for Brain Science and Brain-Inspired Intelligence, Chengdu 611731, China
[e] Montreal Neurological Institute, McGill University, Montreal, Canada
[f] Cuban Neuroscience Center, Ave 25, Esq. 158, #15202, Cubanacan, Playa, Havana, Cuba

ARTICLE INFO

Keywords:
WeBrain
Brainformatics
EEG big data
Distributed cloud computing
Computational ecosystem

ABSTRACT

The current evolution of 'cloud neuroscience' leads to more efforts with the large-scale EEG applications, by using EEG pipelines to handle the rapidly accumulating EEG data. However, there are a few specific cloud platforms that seek to address the cloud computational challenges of EEG big data analysis to benefit the EEG community. In response to the challenges, a WeBrain cloud platform (https://webrain.uestc.edu.cn/) is designed as a web-based brainformatics platform and computational ecosystem to enable large-scale EEG data storage, exploration and analysis using cloud high-performance computing (HPC) facilities. WeBrain connects researchers from different fields to EEG and multimodal tools that have become the norm in the field and the cloud processing power required to handle those large EEG datasets. This platform provides an easy-to-use system for novice users (even no computer programming skills) and provides satisfactory maintainability, sustainability and flexibility for IT administrators and tool developers. A range of resources are also available on https://webrain.uestc.edu.cn/, including documents, manuals, example datasets related to WeBrain, and collected links to open EEG datasets and tools. It is not necessary for users or administrators to install any software or system, and all that is needed is a modern web browser, which reduces the technical expertise required to use or manage WeBrain. The WeBrain platform is sponsored and driven by the China-Canada-Cuba international brain cooperation project (CCC-Axis, http://ccc-axis.org/), and we hope that WeBrain will be a promising cloud brainformatics platform for exploring brain information in large-scale EEG applications in the EEG community.

1. Introduction

The field of brain science research is at the beginning of a new era that combines the burgeoning capabilities of the IT revolution with powerful emerging neuroimaging technologies. First, a number of new or mature technologies, including electroencephalography (EEG), magnetoencephalography (MEG), magnetic resonance imaging (MRI) and functional near-infrared spectroscopy (fNIRS), enable us to noninvasively acquire more varied neuroscientific data at different time and space scales than ever before (Sejnowski et al., 2014). Using these techniques, neuroscience communities around the world, as well as brain projects promoted by national initiatives (e.g., the Human Connectome Project (HCP) with more than 1200 subjects (Glasser et al., 2016) and the European Union Human Brain Project (HBP) with broad cooperation from 117 partner institutions in Europe (Amunts et al., 2016)), produce more complex and extensive neuroscientific datasets, which may imply the arrival of 'big data' age in neurosciences (Sejnowski et al., 2014; Van Horn and Toga, 2014). In addition, analytic strategies are being developed to mine underlying massive information from this wealth of big neuroscientific data, sometimes termed 'big data analysis', including a vast range of univariate and multivariate statistical approaches (Bzdok and Yeo, 2017; Smith and Nichols, 2018). Second, by addressing the large-scale multimodal data acquisition, accumulation and analytics, an increasing number of neuroscience researches are being transformed through the application of high performance computing (HPC), especially cloud computing, which offers abundant resources, such as

* Corresponding author at: The Clinical Hospital of Chengdu Brain Science Institute, MOE Key Lab for Neuroinformation, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu, China.
E-mail address: dyao@uestc.edu.cn (D. Yao).

https://doi.org/10.1016/j.neuroimage.2021.118713.
Received 21 July 2021; Received in revised form 25 October 2021; Accepted 4 November 2021
Available online 17 November 2021.

computational power, memory and storage capabilities, to process and analyze massive multimodal data (Bouchard et al., 2016). Clearly, this era in neurosciences is experiencing a substantial shift from isolated single efforts with limited data, methods and computing resources to a more efforts with larger datasets, more comprehensive methods and cyberinfrastructures. In addition, with recommendation of open science (Koch and Jones, 2016), this transition further generates 'cloud neuroscience' comprising categories of data, infrastructure, apps, algorithms and education to accelerate brain science research (Neuro Cloud Consortium, 2016). Therefore, with the rapid development of IT capabilities and the accumulation of big neuroscience data, the efficient integration of IT infrastructures (e.g., distributed computing recourses) and a wide range of researchers, including neuroscientists (especially the potential researchers who may have few computer programming experience), neuroinformatic method and tool developers and computer engineers and administrators, in a flexible and sustainable brainformatics platform of computational ecosystem, remains a considerable challenge.

Since the first report was published nearly 90 years ago (Berger, 1929), the scalp EEG has been an excellent technique for uncovering brain functions in a wide range of fields, including clinical neurophysiology (Li et al., 2019a; Xu et al., 2014), cognitive neuroscience studies (Enriquez-Geppert et al., 2017; Li et al., 2019b; Tian et al., 2018) and brain-computer interfaces (He et al., 2013; Zhang et al., 2019; Zhao et al., 2020), due to its high temporal resolution (~milliseconds), low cost and noninvasive direct measurement of neuronal activity (Cohen, 2017). Furthermore, in view of the complementarity of the spatiotemporal resolution, the scalp EEG is more valuable when combined with other imaging modalities, such as MRI (Dong et al., 2014, 2015a; Friston et al., 2019; Laufs, 2012). Especially, in clinics, the EEG has been a standard test for diagnosing and characterizing brain diseases, such as epilepsy and strokes, and a basic measurement for the detection of sleep stages (Yamada and Meng, 2012). Over the last decade, the evolution of brain science research has also led to more research with large-scale EEG applications in the EEG community. First, a number of large EEG datasets have been rapidly accumulated and are publicly available worldwide. For example, the TUH-EEG Corpus (https://www.isip.piconepress.com/projects/tuh_eeg/index.shtml) is an open EEG database that contains more than 30,000 clinical EEGs from Temple University Hospital (Obeid and Picone, 2016). The "EEG Motor Movement/Imagery Dataset" (https://archive.physionet.org/pn4/eegmmidb/) contains approximately 1500 EEG recordings from 109 subjects (Goldberger et al., 2000; Schalk et al., 2004). The PREDICT database (http://predict.cs.unm.edu/downloads.php) contains EEG data from 437 psychiatric and neurological patients. The MNI Open iEEG Atlas (https://mni-open-ieegatlas.research.mcgill.ca/) provides intracranial EEG data of 1772 channels from 106 subjects (Frauscher et al., 2018). Furthermore, other open EEG datasets have been released in open repositories such as PhysioNet (https://www.physionet.org/content/) and OpenNEURO (https://openneuro.org/), or are personally collected and shared on the internet (https://sccn.ucsd.edu/~arno/fam2data/publicly_available_EEG_data.html). In addition, our lab also accumulated more than 20,000 clinical EEGs and 2000 normal EEGs to date. However, there are few specific brain information computing platforms for handling those rapidly accumulating EEG big data and supporting large-scale EEG applications. Second, to handle large-scale EEG data, a number of EEG tools and pipelines are expected and have been developed. For example, while most traditional MATLAB-based EEG tools, such as EEGLAB (Delorme and Makeig, 2004), and FieldTrip (Oostenveld et al., 2011), were designed to detail and batch process small scale EEG data for offline personal computer users, many MATLAB-based pipeline tools, including PREP (Bigdely-Shamlo et al., 2015), CTAP (Cowley et al., 2017), HAPPE (Gabard-Durnam et al., 2018) and Automagic (Pedroni et al., 2019), or standards such as EEG Study Schema (ESS, http://www.eegstudy.org/) have been developed to preprocess large-scale EEG raw data and assess the quality of preprocessed EEG data offline. However, most of these tools may be difficult to use for average EEG researchers, especially those with limited programming skills or EEG beginners. These tools often require familiarity with programming skills, the understanding of different tools with different underlying design philosophies and knowledge of configuration parameters and inputs relative to both EEG and toolboxes. Furthermore, for heavier computing loads that need to be solved in HPCs or computing servers, good IT skills, such as those related to the scheduling software environment and policies used at each cloud computing site, are further required, which limit their potential for large-scale EEG studies. In addition, the existence of a variety of options in EEG studies, including varied EEG tools and pipelines, software and operating environments, methods, parameter settings, and even the different researchers (Botvinik-Nezer et al., 2020), may decrease the reproducibility of results without a common computing platform with common pipelines (Nichols et al., 2017). Therefore, these issues represent a sufficient motivation to promote a sustainable EEG cloud computing platform for large-scale EEG applications with common tools and pipelines, which may accelerate brain function discovery via EEG-rich information sources.

Overall, while most current neuroimaging distributed computing platforms were first designed for MRI big data (Makkie et al., 2019; Manjon and Coupe, 2016; Marcus et al., 2007; Sherif et al., 2014) or providing HPC resources and technologies (Amunts et al., 2016; Majumdar et al., 2016) with portal interface plugins (e.g. EEGLAB plugin, nsgportal) (Martinez-Cancino et al., 2021), there are few specific brain information platforms for large-scale EEG big data analysis that would well benefit the EEG community. Therefore, driven by the China-Canada-Cuba international brain cooperation project (CCC-Axis, http://ccc-axis.org/, and the history of CCC-Axis can be seen in http://www.neuro.uestc.edu.cn/neuro/files/CCC_axis.pdf), which takes advantage of the complementary strengths and commonalities from the three partners to forge an integrated computational framework for brain research, a WeBrain platform (https://webrain.uestc.edu.cn/) was designed with following abilities:

(1) Wide computational ecosystem integrating IT infrastructures and researchers with different fields including administrators/IT engineers, brain information method/tool developers and neuroscientists;
(2) Cloud storage and computing with automated accessing and scheduling of computing nodes (e.g., computing servers and HPCs), automated multipoint data movement, and satisfactory isolation and stability of operations from user requests;
(3) Stable and secure service system with lightweight core components, full logging and monitoring of all user actions, low cost deployment and operations, and scalability (without architectural bottlenecks);
(4) Maintainability and sustainability by an IT administrator team;
(5) Standardization, stability and flexibility to develop, integrate and run any tools with any environments;
(6) Convenient web access (no software or system installation required and no computer programming skill required) to realize the end-to-end processing of workflows on the cloud.

In this technology report paper, we will first briefly overview the general scheme of the WeBrain platform, and we subsequently introduce how the above philosophy and guidelines have been implemented in the WeBrain platform. Next, the illustrations, current deployment and

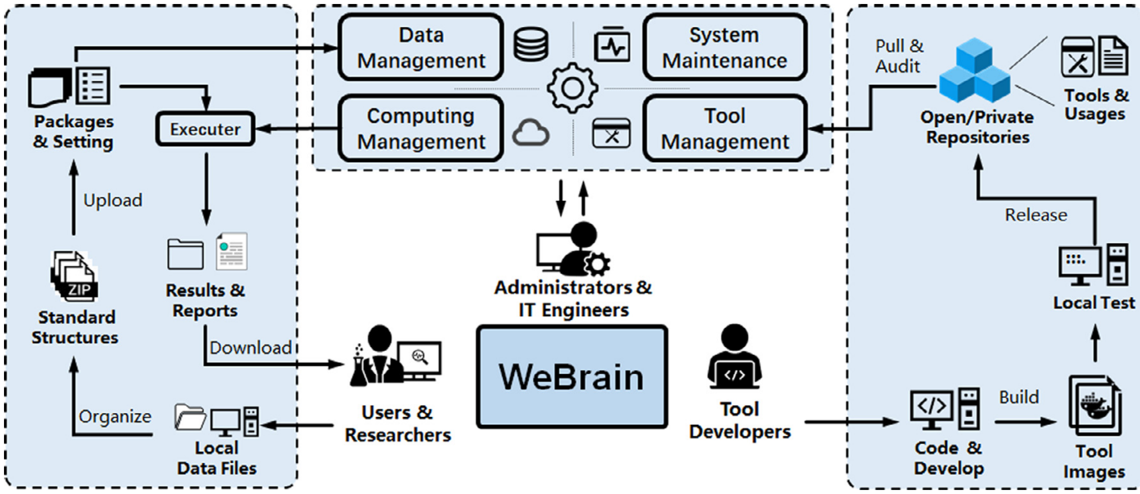

Fig. 1. Computational ecosystem in the WeBrain platform. According to the different demands and experiences of different researchers, three roles including IT engineers/administrators, tool developers and neuroscience researchers are categorized. For IT engineers and administrators, the daily management and maintenance of the WeBrain system, as well as management of computing resources, are easy, flexible and efficient. For tool developers, a flexible API of standardized units (container images) is provided in WeBrain. The developers can develop and build kinds of tool images using the programming language and environment of their choice, test them in the container using personal computers, and then release them on the cloud repositories so that they can be pulled and audited by the administrators. For users such as neuroscience researchers, there is no requirement for computer programming skills or installing any software or system. The workflow of WeBrain contains (1) organizing and uploading local data files to WeBrain, (2) creating and running a computing task, and (3) checking and downloading results to a personal computer. We note that all these roles are associated with WeBrain and could work independently to some degree.

usage of the platform are presented. At last, a discussion and conclusion regarding the platform are provided.

2. Materials and methods

2.1. Computational ecosystem

To efficiently integrate IT infrastructures and researchers in different fields including administrators/IT engineers, brain information method/tool developers and neuroscientists, the WeBrain platform was designed as a sustainable brainformatics platform for a computational ecosystem. In the WeBrain ecosystem (see Fig. 1), three roles are categorized according to the different demands and experiences of different researchers. These roles are associated in the WeBrain ecosystem and could work independently to some degree. These roles further help to form a WeBrain virtual community of brainformatics across worldwide researchers.

2.1.1. Roles in WeBrain

Regarding the roles of administrators and IT engineers, the daily management and maintenance of the WeBrain system and the management of computing nodes are easy, flexible and efficient because of using the microservice architecture and Docker virtualization technique to modularize and containerize the functions and resources of the platform. For administrators, most of the daily operations including registration and privilege management, data files management, tool upgrade and management, monitor of computing tasks and homepage management, are all realized on the WeBrain website. It is not necessary on the command-line interface in the backend server directly. And only the IT team has access to the backend system. Moreover, computing nodes can be conveniently added (authorized) or removed (expired) by the administrators on the website, and these physical computing resources will be virtualized and allocated by the WeBrain system automatically.

Regarding the roles of brain information method/tool developers, to largely decrease the requirements of IT skills and increase the friendliness of the WeBrain platform, a flexible and conceptual application programming interface (API) of standardized units (container images) is defined by WeBrain. The API is defined as a function with inputs and outputs, which is in line with most of tool usages on command window. The first 3 inputs (“Input”, “Output” and “combs_project_id”) of tools are fixed and required (The ‘Input’ is a string of zip data file direction, the ‘Output’ is a string of output direction and the ‘combs_project_id’ is a task ID, respectively), and following parameters (strings) can be defined by developers (more details see the document on https://webrain.uestc.edu.cn/documentation.html, “*How to Develop Standard WeBrain Tools*”). Considering the consistency and update of tools developed by different languages (e.g., C/MATLAB/Python/JAVA) in different operating environments (Windows/Linux/Mac), tools in WeBrain are standardized units containerized by the virtualization technique such as Docker (https://www.docker.com/). It is suggested that one generates a standardized unit (container image) that includes everything needed to run the application (e.g., operating system, environments and codes) for development and shipment using the container technique; therefore, tools can be integrated and called in WeBrain, no matter what programming languages and environments are used. Therefore, developers can develop all kinds of tools with the WeBrain API definition using the programming language and environment of their choice, test them in the container using their own computers, and then release them (as well as tool instructions and examples) on a cloud repository so that they can be pulled by the WeBrain system. Noting that, for tool developers, there are few requirements of IT skills and experiences such as k8s configuration, RESTful web service, dashboard implementation and computing resource scheduling etc. Tool images will be pulled and audited by the WeBrain administrator first, and then be added and released on the WeBrain platform.

Regarding the role of neuroscience researchers (registered users), WeBrain provides the good user experience of working on the cloud. In the WeBrain platform, there is no requirement for computer programming skills or installing any software or system for users. All that is needed is a modern web browser, which simultaneously reduces the technical expertise required to use distributed cloud computing and the kinds of basic and specific neuroscience tools. The workflow of WeBrain contains (1) organizing and uploading local data files to WeBrain, (2) creating and running a computing task, and (3) checking and downloading results to a personal computer. All users will be benefit from these tool and cloud computing resources that are easy to use.

2.1.2. WeBrain community

The WeBrain platform is sponsored and driven by the CCC-Axis. WeBrain is guided by the CBRAIN (http://mcin.ca/technology/cbrain/) (Sherif et al., 2014) team of the Canadian partners at the Montreal Neurological Institute of McGill University. And, it may be a natural companion piece of CBRAIN to go together to promote cooperation researches under CCC-Axis. Meanwhile, due to the three parties' decades of accumulated EEG experience, specific/basic methods and tools developed by three parties will be integrated in the both platforms, and the WeBrain platform may be a community with natural EEG specialties of data management and processing. Thereby, an active and engaged virtual community, named the International Virtual Community of Brainformatics (IVCB), is being formed to set the scene for multinational initiatives and cooperation in brain research. And a WeBrain forum is being developed to realize and promote the virtual community for WeBrain users. In this virtual community, researchers in different fields can exchange thoughts and share experiences on EEG researches. Currently, the WeBrain platform is continuing based on the cooperation with members of the CCC-Axis (http://ccc-axis.org/) and Joint China Cuba Joint China Cuba Lab for Frontiers Research in Translational Neurotechnology (https://www.neuroinformatics-collaboratory.org/). And, WeBrain is participating in two projects, the EEGManyPipelines project (https://www.eegmanypipelines.org/) and the carbon footprint quantification project of the OHBM Neuroimaging Research Pipelines Workgroup (https://neuropipelines.github.io/index). In addition, 1st WeBrain workshop (more than 200 participants) and 1st WeBrain EEG training course (about 40 participants) were successfully held in Chengdu. In the future, the WeBrain training course is planned to be held twice a year (online and offline).

2.2. WeBrain architecture overview

The architecture of WeBrain is composed of three main layers (See Fig. 2). (i) The user layer is considerable service consumption (i.e. requests from users) through web browsers. (ii) The platform layer hosts the WeBrain service that is responsible for all requests from the top user layer, and it realizes users, data and tool management and computing resource scheduling. (iii) The cloud resource layer represents the wide cloud resources of tool repositories, data storage and computing nodes.

2.3. Distributed cloud computing

The first challenge faced by WeBrain was how to manage and schedule computer resources (e.g., computing servers and HPCs). Creating standard and virtual units (e.g., containers) of these computing resources and then developing a centralized system to manage and schedule these virtualized units would be a flexible and stable way to overcome this challenge. In the WeBrain system, the process virtualization technique such as Docker was used to orchestrate and manage application containers. All Docker containers are standard, isolated, portable and lightweight units that run on the Docker Engine (https://www.docker.com/) (Felter et al., 2015; Merkel, 2014). Then, a subsystem for automated COntainer Management and Batch Schedule (COMBS) was developed to enable computing workloads on cloud infrastructures using Kubernetes (k8s, https://kubernetes.io/) (Netto et al., 2017) container orchestration mechanisms. In addition, the above strategy and realization are invisible to users.

Considering the master site is the entity hosting all WeBrain service, and computing sites are clusters with private storage and computing nodes, COMBS is designed and runs on master site for the management and schedule of computing tasks and containerized computing resources. When receiving a request containing information about the user ID, computing task ID, tools and parameters to use from users, COMBS can then automatically package and push the job request to a computing site which user belongs (neuroimaging data files are not transmitted among computing sites). And then, a short-lived one-off task (i.e. an applicable computing container) according to the loads and types (CPU or GPU) of computing nodes on the computing site will be created. Because the data file directory on the data storage nodes is mounted into the container, the data files required for analysis are directly loaded and calculated on the computing nodes, and results will be written back to mounted storage nodes. Once the analysis is done, the COMBS initiates transfers of the finished job information to the master site. Those temporary data and results in the memory on the computing nodes will be cleaned. COMBS also performs activities including configuring computing nodes (e.g., adding or removing computing nodes and the type of computing nodes), containerization (e.g., the number of containers per node, the number of CPUs and memory per container), monitoring containers of computing nodes, recovering failed computing containers, and queueing jobs when computing resources are completely consumed.

2.4. Maintainable and scalable microservice architecture

The concept of a ‘microservice architecture’ was first proposed and discussed by Fowler and James in 2011 at a workshop of software architecture and formed (Fowler and Lewis, 2014). Compared with a traditional service architecture, the microservice architecture is a new approach to develop applications/service systems by composing of a suite of small and independent services, with each running its own processes and communicating via lightweight mechanisms (Dragoni et al., 2017; Fowler and Lewis, 2014). These small services are built around system functionalities and could be independently developed and deployed. The microservice architecture delivers many benefits in terms of its small size, high flexibility, maintainability and scalability and is becoming a new trend in software architecture, both in academia and the industrial field. The WeBrain system was therefore developed based on the above-mentioned microservice architecture. Basic functionalities of WeBrain, including the foregrounding web server, management and scheduling of computing tasks and resources, tools, data and user management, are functionally decomposed into a set of small services. Each small service in the WeBrain system is operationally independent from other service modules and communicates with other services through defined interfaces. Furthermore, because the adopted microservice architecture naturally makes containerization applicable to these microservices, new service modules of the WeBrain system can be directly developed, tested and maintained in isolation with respect to the rest of the WeBrain system by IT engineers using optimal languages and frameworks. Furthermore, changing or scaling a service module of the WeBrain system requires rebooting or updating the microservices of that module only and not the whole system. In addition, the foregrounding service module realizes the convenient web access to visit the WeBrain platform for users (no software or system installation required and no computer programming skill required).

2.5. Privacy and security

WeBrain was designed with the well-practice privacy and security strategies at the forefront to enable secure access to neuroimaging data and results, protect the safety of database information, and communicate between the master site and computing sites. In the WeBrain platform, users only have access to data and result files they have added or calculated through their own account, and user isolation ensures that the users can not access any information and data of the other users of the platform. The current way to download or share data and result files is using a temporary download link (expires in 30 days), which is generated and sent to the registered email address by the WeBrain system. This approach means that only the users can decide to whether and when to share their data or results by emailing temporary download links, and this largely reduces the risk of data theft. Furthermore, snapshot technology was used to regularly back up system database records to ensure that the WeBrain system can be rebuilt at neatest point in

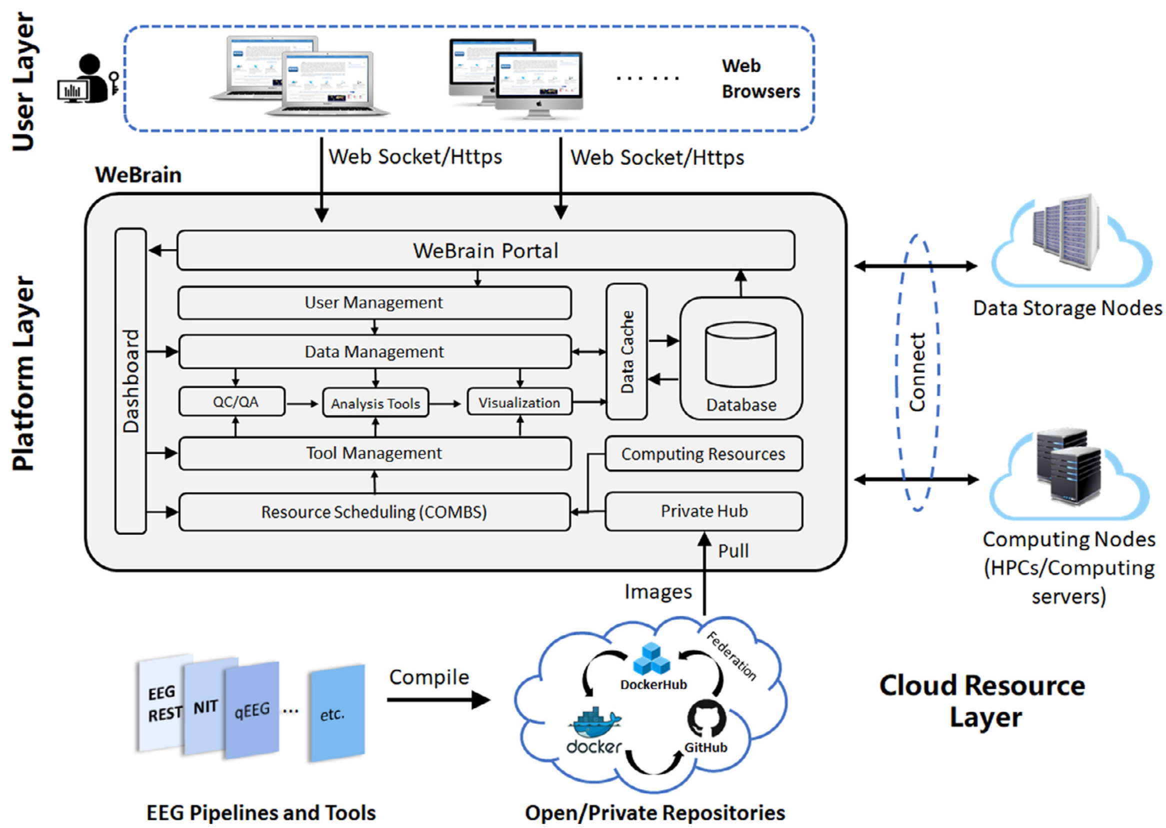

Fig. 2. Architecture of WeBrain. The top layer represents the service consumption of web browser users. The central platform layer hosts the WeBrain service that is responsible for requests from the top user layer; realizes user, data and tool management; and conducts computing resource scheduling. Noting that, Redis (https://redis.com/) is used as a caching solution with cache-aside and write-back strategies in WeBrain. The lower right cloud resource layer represents wide cloud resources of tool repositories, data storage and computing nodes. And, WeBrain is flexible and can expand and reduce resources according to the actual demand.

time. Communications in WeBrain (e.g., visiting website, job requests and system database information) is masked and encrypted based on Hyper Text Transfer Protocol over Secure Socket Layer (HTTPS), which is a more secure protocol for authentication of the accessed website, and protection of the privacy and integrity of the exchanged data while in transit.

2.6. Current tools in WeBrain

Many of the basic or specific EEG tools are integrated in the WeBrain platform, since they are commonly used or computationally expensive and generally complex to use for the novice user (Table 1). Currently, for EEG preprocessing, WeBrain includes (1) an EEG re-referencing tool for the reference electrode standardization technique (REST) developed by our group (Dong et al., 2017; Yao, 2001; Yao et al., 2005), (2) a stable quality assessment (QA) tool for large-scale continuous EEG raw data, and 3) standardized and specific EEG preprocessing pipelines of large-scale continuous EEG raw data to remove various artifacts. For further EEG analysis, WeBrain includes (1) power spectrum analysis based on time-frequency analysis adapted from EEGLAB (Delorme and Makeig, 2004), (2) EEG network and topology analysis based on graph theory (partially adapted from the brain connectivity toolbox available on https://sites.google.com/site/bctnet/ (Rubinov and Sporns, 2010)), and (3) event-related potential (ERP) analysis at the scalp level. For EEG source imaging analysis, WeBrain includes (1) the generation of the conduction model of a real head based on the boundary element method (BEM) using standard MRI T1 images (adapted from FieldTrip (Oostenveld et al., 2011)), and (2) scalp EEG source imaging based on a forward model (3-concentric sphere head model or real head model) and inverse method such as sLORETA (Dale et al., 2000; Pascual-Marqui, 2002). Especially, the EEG REST is a toolbox for a reference electrode standardization technique (Dong et al., 2017; Yao, 2001; Yao et al., 2005) for translating EEG data with a reference to any a physical point on brain/body surface or the postprocessed data referenced at average or linked ears to a new dataset with a reference at infinity where the potential is zero/constant. Meanwhile, the REST has been listed in the new guidelines of the International Federation

Table 1
EEG tools/methods and file formats currently supported by the WeBrain platform (see https://webrain.uestc.edu.cn/documentation.html).

	Supported tools/file formats
EEG tools/methods	REST
	Bad Block Marking
	ICA Running
	Quality Assessment
	Preprocessing of Continuous EEG Raw Data
	Power Spectrum Analysis
	Event-Related Potential Analysis
	EEG Network Calculation
	Network Topology Analysis
	Generation of Real Head Conduction Model (BEM)
	EEG Source Imaging
EEG file formats (Manufacturer/file format)	ASCII/Float file (*.txt)
	MATLAB (*.mat/*.dat)
	EEGLAB (*.set/*.fdt)
	Curry 6/7 (*.dat/*.dap/*.rs3)
	Curry 8 (*.cdt)
	BrainProducts/Brain Vision (*.vhdr/*.vmrk/*.dat)
	NeuroScan (*.cnt/*.eeg)
	Biosemi/European Data Format (*.bdf/*.edf)
	BIOSIG (*.edf/*.edf+/*.gdf/*.bdf)

of Clinical Neurophysiology (IFCN) for EEG analysis (Babiloni et al., 2020). The EEG preprocessing tool is a stable and specific pipeline to remove artifacts of continuous EEG raw data. The tool integrates different kinds of methods, including the quality assessment of raw data, passband and notch filtering, artifact removal using electronystagmogram (EOG) regression and residual artifact removal methods (ICA-based Multiple Artifact Rejection Algorithm (MARA) (Winkler et al., 2011), ICA-based ADJUST (Mognon et al., 2011), robust PCA (Lin et al., 2010) or artifact subspace reconstruction (ASR) method (Mullen et al., 2013, 2015)), bad channel interpolation (spherical spline interpolation (Perrin et al., 1989) and REST re-referencing or REST-based interpolation method (Dong et al., 2021)), quality assessment of preprocessed data and marking residual bad blocks. Then, the clean EEG data with REST reference are finally obtained. The source imaging tool is a stable and convenient pipeline to estimate source signals of scalp EEG/ERP data based on a forward model and inverse method. The approach consists of checking items including data, channel locations and sampling rate, calculating lead-field matrix based on a 3-concentric sphere head model or real head model, filtering and extracting interested epochs and estimating source signals using sLORETA inverse method (Dale et al., 2000; Pascual-Marqui, 2002). In addition, if needed, it could automatically match source signals to the Automated Anatomical Labeling (AAL) brain template (Tzourio-Mazoyer et al., 2002) to obtain the averaged source signals of brain regions. These averaged source signals could be further used for EEG network and topology analyses. More details about these tools, their usages and contributors can be seen in the tool instructions, which can be downloaded from the WeBrain website (https://webrain.uestc.edu.cn/documentation.html). In addition, other resources and personal computer tools including EEG-fMRI multimodal fusion toolbox, the Neuroscience Information Toolbox (NIT, https://www.neuro.uestc.edu.cn/NIT.html) (Dong et al., 2018), the emiCCA toolbox (Dong et al., 2015b), the EEG network analysis (ENA) toolbox, and WeBrain example datasets and collected links to open EEG datasets are also available on the WeBrain website (https://webrain.uestc.edu.cn/resources.html).

3. Results

3.1. Illustrations of usage

Users can visit, register personal accounts and then login to WeBrain on the website (the WeBrain portal is here: https://webrain.uestc.edu.cn/index.html). The usage of WeBrain is all realized on the website, and the main components of the user environment are shown in Figs. 3–7, namely, the “My Monitor”, “My Project”, “My File”, “My Statistic” and “Account Settings”. The “My Monitor” module (Fig. 4) shows the state of all computing tasks with different colors (green: succeed; red: failed; blue: terminated and gray: other). The basic information of computing tasks including the task name, used tool, selected data and creator, is shown in the box. The “My Project” module (Fig. 5) shows the list of computing jobs from user operations. Users can create a new computing task by clicking the button “+new task”, filling in the task name and description, selecting a tool, providing the parameters, selecting data files and clicking the “execute” button to run the task (or the “save” button to save the task). Users can also edit the computing task (undone task only), see task details and delete unwanted tasks (waiting, failed or undone tasks can be deleted only) by clicking the related buttons. All data transfers, container setup, scheduler interactions and monitoring are handled behind the scenes by the WeBrain system. The “My File” module (Fig. 6) shows the list of all data files (blue) uploaded by users or result files (green) calculated in the WeBrain platform. Users can filter, manage, delete and download data files through a graphical user interface. Note that users only have access to the data in their account, and they can share/download the results or data files through a temporary download link generated by the WeBrain system; thus, only the users can decide whether and how to share the files. The “My Statistic” module (Fig. 7) shows the records (e.g., number of computing tasks, number of uploaded files, and file sizes) of the usage of a current user during a period of time. At last, users can change their profiles and passwords through the “Account Settings” module. More details can be seen in the WeBrain manual for users (https://webrain.uestc.edu.cn/documentation.html).

3.2. Example and reliability of WeBrain

To explore the reliability of the computations performed via WeBrain via two runs performed on the two personal computers, an example EEG dataset during the resting state (eyes closed) was used. A total of 40 healthy subjects (mean age ± standard deviation = 24 years ± 1.6 years; 31 males and 9 females) were recruited. Written informed consent in line with the Declaration of Helsinki were obtained before the resting state EEG recording. The experiment was also approved by the local ethics committee of the University of Electronic Science and Technology of China (UESTC). The raw data were first preprocessed using the WeBrain EEG preprocessing tool “WB_EEG_prepro” and then utilized to calculate the relative power indices in various frequency bands using the WeBrain tool “WB_EEG_CalcPower”. The broad bands were defined by the guidelines (Babiloni et al., 2020; Jobert et al., 2013; Malver et al., 2014). The same deterministic processing steps (calculating power indices) were also performed on two personal computers using MATLAB codes adapted from EEGLAB. At last, the intraclass correlation coefficient (ICC) of the relative power indices was calculated across WeBrain and local personal computer runs. More details of the experimental paradigm and the EEG recording and processing steps applied to the EEG raw data can be seen in the supplementary materials and the article (Li et al., 2015).

Fig. 8 shows a preprocessing result of an example EEG raw data from a subject using the WeBrain platform. Artifacts were well removed by the preprocessing tool and a bad channel of AF7 was also automatically detected and reconstructed. The results of the power analysis are shown in Fig. 9, in which the relative power indices in different frequency bands are shown in the first 3 rows. As we can see, the dominant delta relative power was found to be in the prefrontal area, and the theta relative power was at fronto-central area. The alpha1 relative power was distributed at the occipital and posterior area with a fronto-central extension, and the alpha2 relative power was at the occipital and posterior area. The dominant beta1 and beta2 relative power was over the parietal and temporal area. The beta3 and gamma relative power were limited mainly at the temporal and occipital area. Fig. 9 also shows the ICC maps across the 3 runs (the ICC values are almost 1). The identical similarity and reproducibility of the relative power maps demonstrate the high level of reliability of WeBrain tool output across runs.

3.3. Current deployment and use

The first WeBrain infrastructure (UESTC master site) is currently physically located in the University of Electronic Science and Technology of China, Chengdu, China. The current production deployment of WeBrain consists of approximately 464 TB storage, 9 CPU nodes (~31.15 teraFLOPS) and 2 GPU nodes (~15.33 teraFLOPS), totaling more than 600 CPU cores, 25,700 CUDA cores and 2.5 TB of memory. For a given job, a container will be newly built and the current limits of CPU and memory of each container are “minimal 0.1 CPU cores (1 CPU = 1000 milliCPUs) and 100 MB memory” and “maximal 2 CPU cores, 2GB memory”, and the current maximal running time of a job is 7 days. Meanwhile, considering to increase the security and stability of the system, the limitation of the CPU loads is first set as 90% (about 500 CPU cores), and the maximum of tasks/jobs running simultaneously is set as 3500, as 70% of theoretical maximum value (5000 tasks/jobs = 500 CPU cores/0.1 CPU cores). The subsystem COMBS in WeBrain could automatically optimizes and allocates an applicable computing container

Fig. 3. The homepage and portal of the WeBrain platform (https://webrain.uestc.edu.cn/).

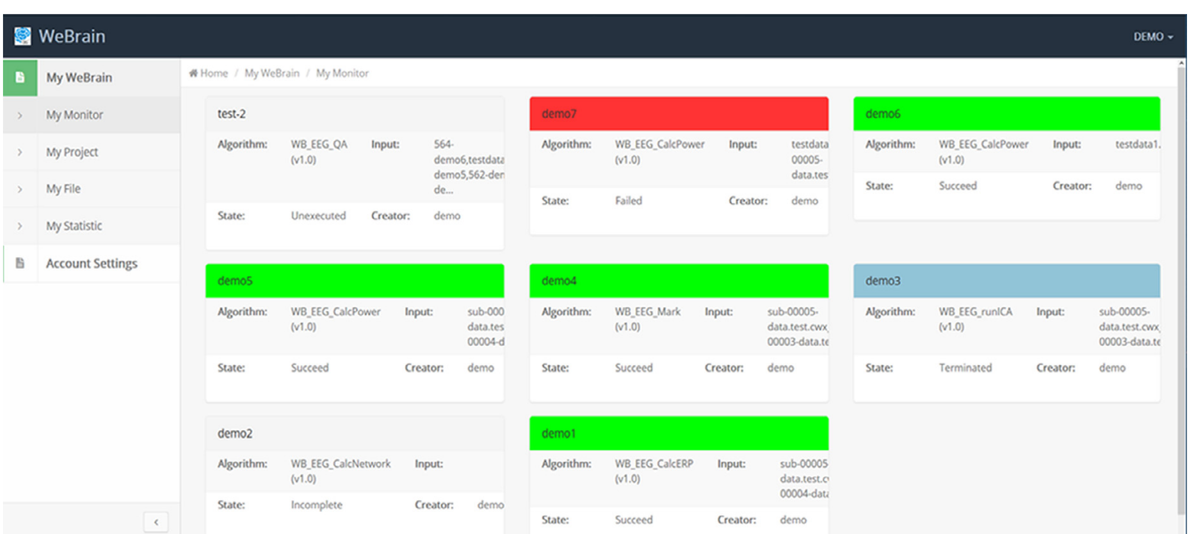

Fig. 4. "My Monitor" module. It shows the basic information (task name, used tool, selected data and creator) of all computing tasks in the boxes with different colors (green: succeed; red: failed; blue: terminated and gray: other) (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

according to loads of computing resources. And, if the number of running jobs exceeds the maximum value, the new computing requests will be queued up. In addition, a saturating test using an EEG tool (EEG preprocessing) was conducted on the WeBrain system. Results showed that the WeBrain system performed well to automatically and dynamically manage and schedule computing resource allocation without crashing (Fig. 10).

To our knowledge, currently, at least 51 countries/areas and 6700 visitors had already visited the WeBrain website (https://webrain.uestc.edu.cn/) a total of 19,000 times, and approximately 230 users from 45 institutes in 25 cities around the world have registered with WeBrain. In addition, materials including the manuals, example data and collected links of open datasets are also provided on the WeBrain website.

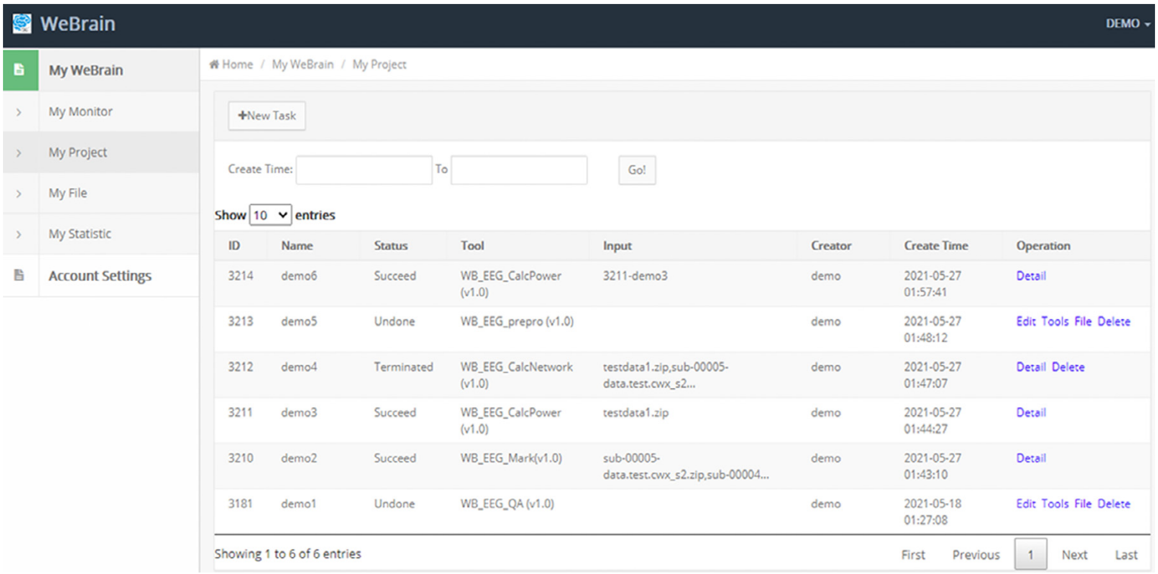

Fig. 5. "My Project" module. It lists all computing tasks from user operations. Users can easily create a new computing task by clicking the "+new task" button, providing the task name and description, selecting a tool, providing parameters, selecting data files and then clicking the "execute" button to run the task.

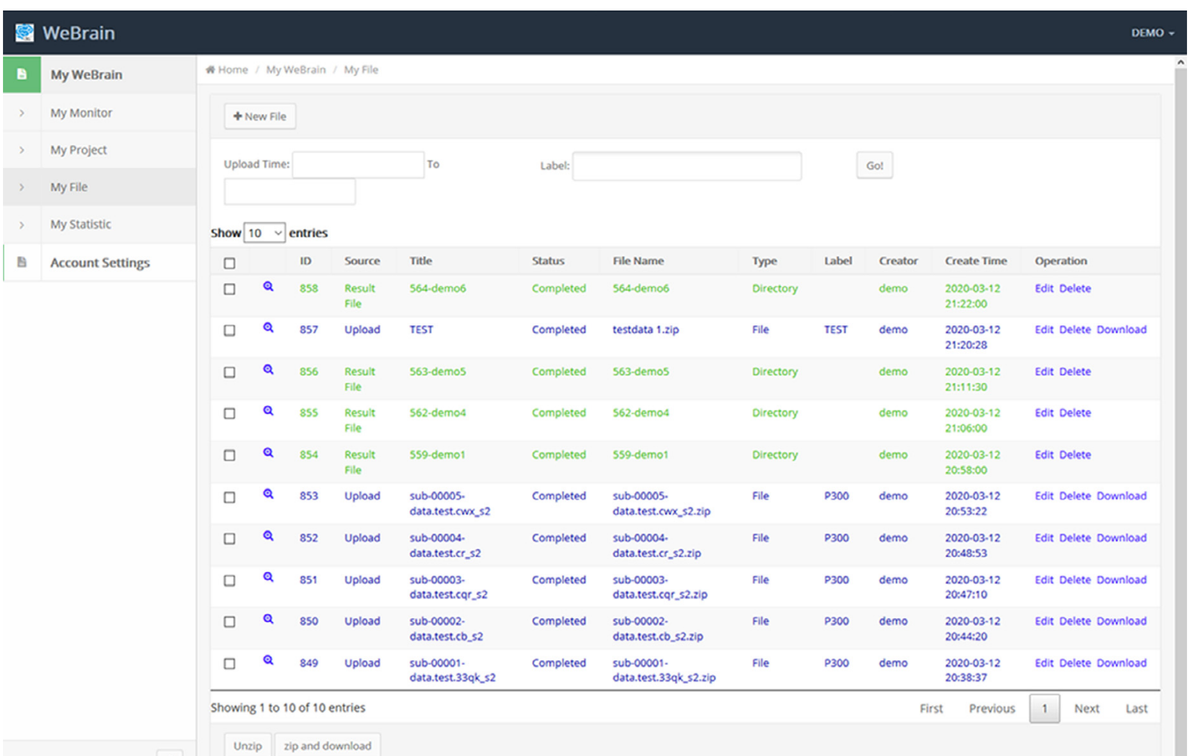

Fig. 6. "My File" module. It lists all data files (blue) uploaded by users or result files (green) calculated in the WeBrain platform. Users can filter, manage, delete and download data files through the web interface (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

4. Discussion and conclusion

With a trend of the 'cloud neuroscience' (Neuro Cloud Consortium, 2016), the WeBrain platform was developed for large-scale EEG applications in the EEG community. The key goals of the WeBrain platform are the following: (1) to provide an end-to-end cloud solution for EEG big data analytics with common tools and pipelines, (2) to construct a sustainable and wide computational ecosystem of brainformatics to bring researchers with different fields and experiences together to maximize their impact, and (3) to increase the computational reliability and reproducibility by using same methods with same codes running on the same environment.

In this paper, we introduced our endeavors to address each of the above goals by reporting the scheme, features and demo usage of the WeBrain platform. First, with challenges in groups of cloud storage/computing and core service systems, the WeBrain system was introduced to form an efficient and specific brainformatics platform for state-of-the-art end-to-end cloud solution for EEG big data analysis. All phys-

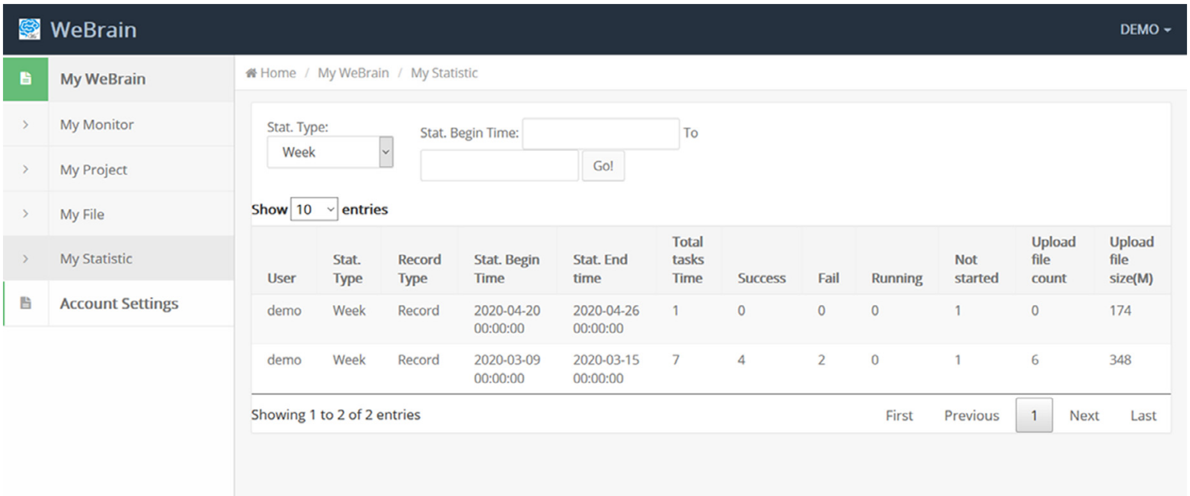

Fig. 7. "My Statistic" module. It shows the records of the usage of the current user during a period of time. Basic information including the number of computing tasks, number of uploaded files and file sizes is displayed.

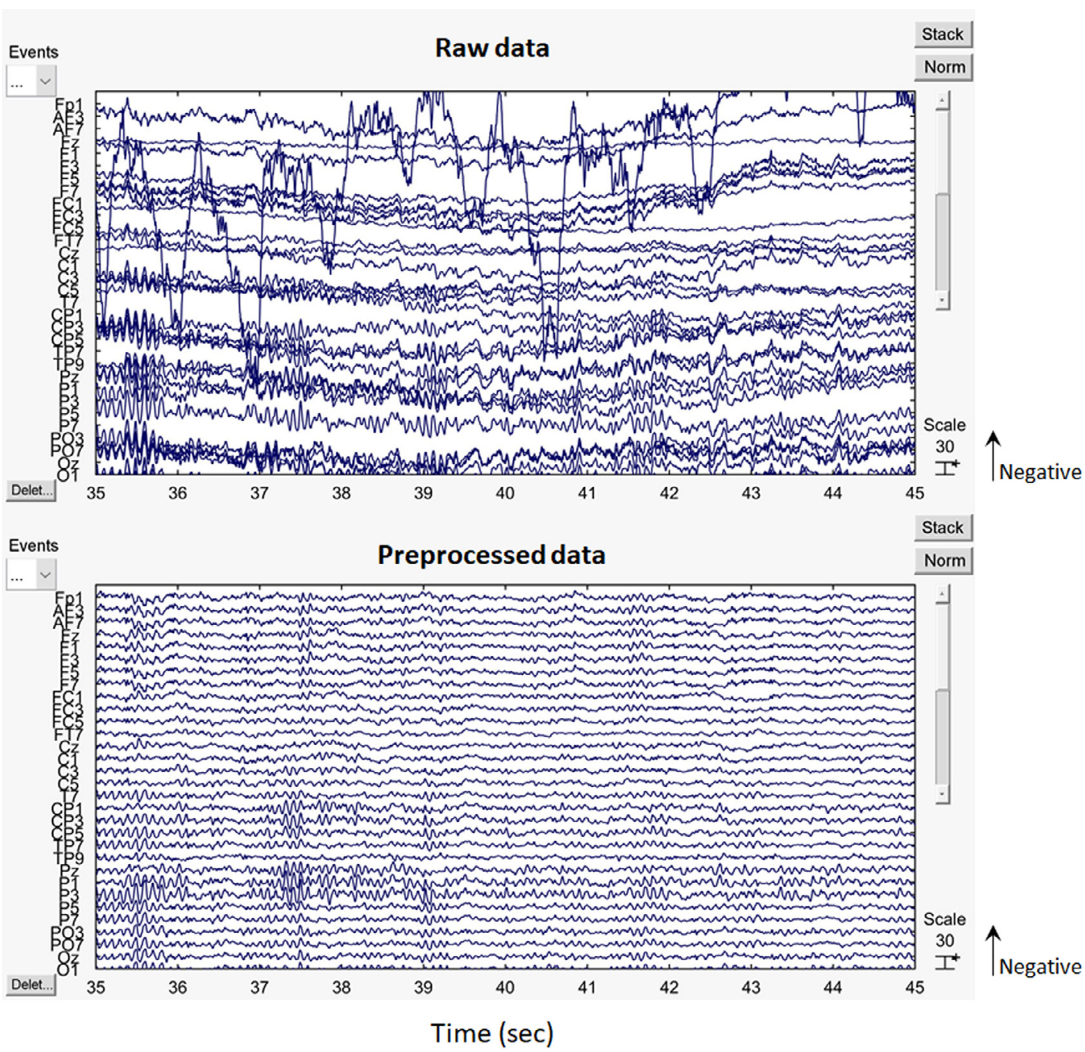

Fig. 8. Preprocessing result of example EEG raw data from a subject using the WeBrain platform. The upper figure shows the EEG waves of raw data with the original recording reference. The bottom figure shows the EEG waves of clean data with the REST reference. The raw data were preprocessed by the WeBrain platform using the "WB_EEG_prepro" EEG preprocessing tool.

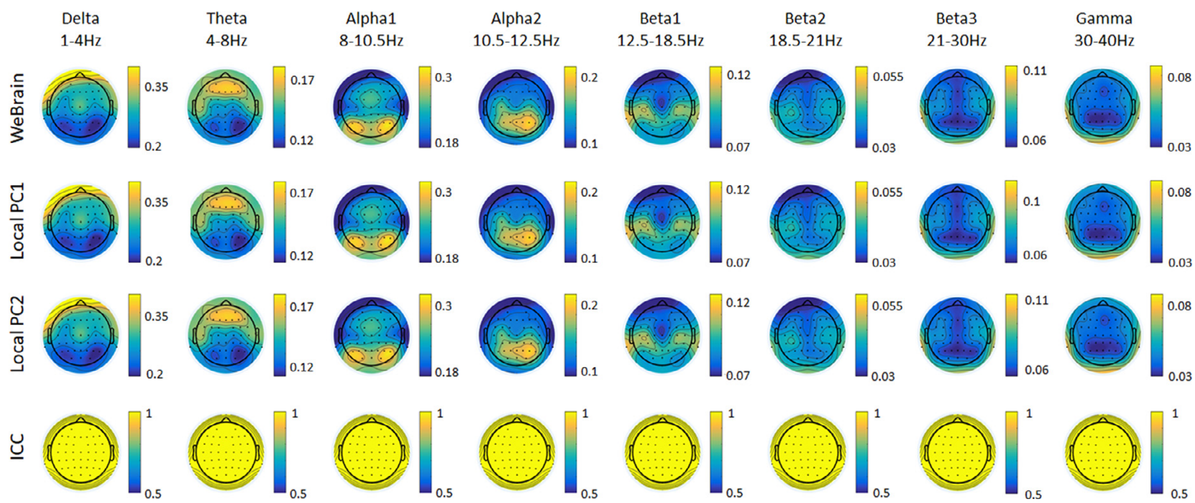

Fig. 9. Example results of WeBrain for calculating the relative power indices. The mean relative power indices across subjects in different frequency bands calculated by WeBrain and local personal computers 1 and 2 (rows 1–3) are shown in the first 3 rows, respectively. The last row shows the ICC values of the relative power indices across WeBrain and local personal computer runs.

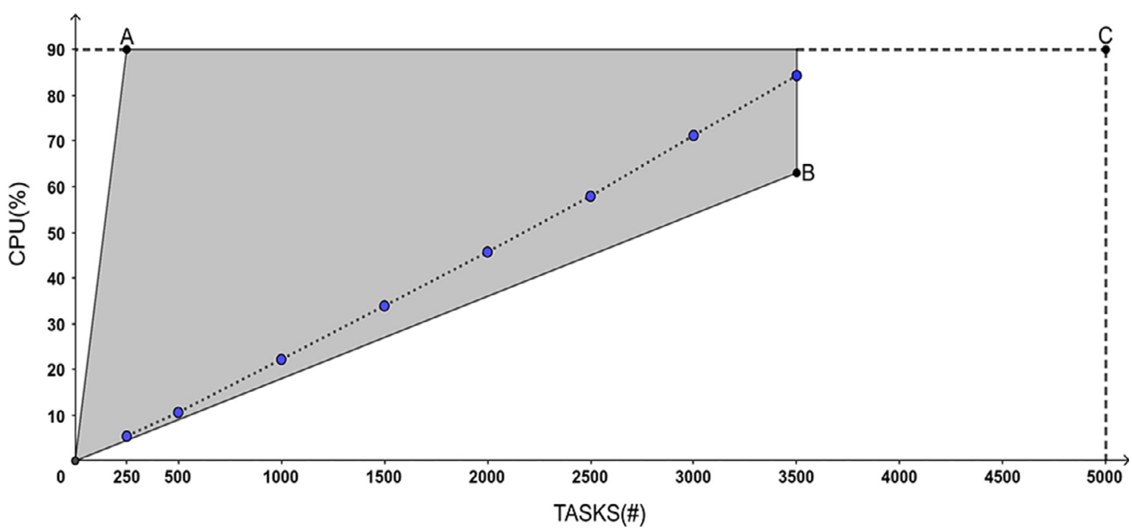

Fig. 10. Results of saturating test on WeBrain. The gray region represents the possible CPU loads under a number of running tasks/jobs. Point A represents 90% CPU load with 250 tasks (maximal 2 CPUs), point B represents ~63.6% CPU load with 3500 tasks (minimal 0.1 CPUs), and point C is theoretical maximum value (90% CPU load with 5000 tasks). The black dotted line represents the system loads correspond to running a number of tasks/jobs (250, 500, 1000, 1500, 2000, 2500, 3000 and 3500 tasks/jobs) using an EEG tool (EEG preprocessing with default parameter settings).

ical computing nodes (e.g., computing servers and HPCs), as well as the common tools and pipelines, were virtualized and run on the Docker Engine (Felter et al., 2015; Merkel, 2014). Then, using Kubernetes (Netto et al., 2017) container orchestration mechanisms, the COMBS subsystem was developed and implemented to enable distributed computing on cloud infrastructures. Meanwhile, the WeBrain system was developed by using a microservice architecture, which is a new trend to develop service systems with the benefits of small size, high flexibility, maintainability and scalability (Dragoni et al., 2017; Fowler and Lewis, 2014). It is noted that although the current focus of WeBrain is primarily on EEG (as well as EEG-fMRI) analyses, it is could be a more generic platform that can accept data and analysis tools from any discipline. Second, considering the usage scenarios for workers/researchers with different amounts of experience, we categorized the current WeBrain computational ecosystem into the 3 roles of administrators and IT engineers, tool developers and users (neuroscience researchers). These roles could work independently to some extent and associated with WeBrain. Then, the system can further help to form a wider virtual community of brainformatics across researchers with different experiences and fields. In the results, we also described the interfaces and primary usage of WeBrain for registered users. For users, it is noted that the data upload and download times may vary depending on the bandwidth and then-existing traffic load of each link in the user connection to WeBrain. Third, we tested the reproducibility of the computations performed via the WeBrain platform on the example EEG dataset. The results show that the EEG relative power indices in different frequency bands during eye-closed resting-state are well replicated and consistent with previous studies (Chen et al., 2008; Yao et al., 2005, 2019). Furthermore, the ICC results show that all EEG relative power indices are identical across the WeBrain platform and local personal computer runs. Because of using Docker container technique to reproduce the runtime environment of neuroimaging tools, the WeBrain framework with neuroimaging pipelines may promote the desired reproducibility for large-scale EEG analysis.

More efforts with the WeBrain platform will be made in the future. First, our current computational resources are limited, and the number of computing nodes will be gradually expanded by adding new computers and HPCs locally in the future. Meanwhile, considering potential users who may be bound by privacy regulations about transferring datasets, more computing sites are welcome to be added in the WeBrain computational ecosystem in the future. And these users could be registered to a computing site which is legally allowed to store and compute certain datasets. Noting that, given the scalability and modularity of the WeBrain system, the computational resources could be flexibly expanded and reduced according to the actual demand by using external cloud-based nodes. Second, priorities for the platform upgrades include the further development of new functionalities and modules, including visualization and knowledge sharing and the refinement of the web API, to interconnect with the companion platform of CBRAIN (Sherif et al., 2014). Third, more common and specific tools/methods on EEG, as well as other modality data, are encouraged to be integrated in the WeBrain system. And, as an extension to the Brain Imaging Data Structure (BIDS) specification for EEG, the EEG-BIDS (Pernet et al., 2019) is supported in WeBrain, and the BIDS-apps are also planned to be integrated in WeBrain in the future. Furthermore, a general solution for EEG big data analysis combined with deep learning techniques using GPU based parallel computing engines to develop new processing tools for large-scale EEG applications is planned. Fourth, to further improve the platform, the full disk encryption of the data will be considered in the next version of the WeBrain system. At last, currently, WeBrain is free for all users (we limit the storage space to 100GB per user), and any users are encouraged to report suggestions, problems and bugs about WeBrain via email to the WeBrain team (Lidong@uestc.edu.cn).

In a conclusion, we have presented a state-of-the-art web-based brainformatics platform as a computational ecosystem, named WeBrain, that enables end-to-end large-scale EEG data analysis on the cloud. We hope that the online nature of the WeBrain platform will facilitate the access of all users with different experiences and fields to WeBrain and will make their tools or EEG data analysis simpler and more efficient, especially for large-scale EEG applications.

Data and code availability statements

The datasets used in this study are available on request to the corresponding author. Materials including demo datasets, WeBrain manual, tool instruction and collected links (open EEG datasets or free EEG tools) etc. can be seen on WeBrain website (https://webrain.uestc.edu.cn/). The compiled code of EEG tools in WeBrain can be downloaded from the link (https://github.com/webrain2018/EEG-Tools). If necessary, the original codes of the EEG tools are available on request to the corresponding author.

CRediT authorship contribution statement

Li Dong: Software, Formal analysis, Writing – original draft. **Jianfu Li:** Software, Methodology, Validation. **Qiunan Zou:** Software, Visualization, Validation. **Yufan Zhang:** Software, Validation. **Lingling Zhao:** Software, Validation. **Xin Wen:** Software, Validation. **Jinnan Gong:** Software, Validation. **Fali Li:** Resources, Validation. **Tiejun Liu:** Writing – review & editing. **Alan C. Evans:** Writing – review & editing. **Pedro A. Valdes-Sosa:** Writing – review & editing. **Dezhong Yao:** Conceptualization, Supervision, Writing – review & editing.

Declaration of Competing Interest

All authors have no conflict of interest to disclose.

Credit authorship contribution statement

Li Dong: Software, Formal analysis, Writing – original draft. **Jianfu Li:** Software, Methodology, Validation. **Qiunan Zou:** Software, Visualization, Validation. **Yufan Zhang:** Software, Validation. **Lingling Zhao:** Software, Validation. **Xin Wen:** Software, Validation. **Jinnan Gong:** Software, Validation. **Fali Li:** Resources, Validation. **Tiejun Liu:** Writing – review & editing. **Alan C. Evans:** Writing – review & editing. **Pedro A. Valdes-Sosa:** Writing – review & editing. **Dezhong Yao:** Conceptualization, Supervision, Writing – review & editing.

Acknowledgments

The present study was financially supported by the National Natural Science Foundation of China (No. 81861128001, 81701778), the '111′ project (B12027), the Sichuan Science and Technology Program (2018JZ0073) and the CAMS Innovation Fund for Medical Sciences (2019-I2M-5–039).

References

Amunts, K., Ebell, C., Muller, J., Telefont, M., Knoll, A., Lippert, T., 2016. The human brain project: creating a European research infrastructure to decode the human brain. Neuron 92, 574–581.

Babiloni, C., Barry, R.J., Basar, E., Blinowska, K.J., Cichocki, A., Drinkenburg, W., Klimesch, W., Knight, R.T., Lopes da Silva, F., Nunez, P., Oostenveld, R., Jeong, J., Pascual-Marqui, R., Valdes-Sosa, P., Hallett, M., 2020. International Federation of Clinical Neurophysiology (IFCN) - EEG research workgroup: recommendations on frequency and topographic analysis of resting state EEG rhythms. Part 1: applications in clinical research studies. Clin. Neurophysiol. 131, 285–307.

Berger, H., 1929. Über das elektrenkephalogramm des menschen. Eur. Arch. Psychiatry Clin. Neurosci. 87, 527–570.

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.M., Robbins, K.A., 2015. The PREP pipeline: standardized preprocessing for large-scale EEG analysis. Front. Neuroinform. 9, 16.

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., Benoit, R.G., Berkers, R., Bhanji, J.P., Biswal, B.B., Bobadilla-Suarez, S., Bortolini, T., Bottenhorn, K.L., Bowring, A., Braem, S., Brooks, H.R., Brudner, E.G., Calderon, C.B., Camilleri, J.A., Castrellon, J.J., Cecchetti, L., Cieslik, E.C., Cole, Z.J., Collignon, O., Cox, R.W., Cunningham, W.A., Czoschke, S., Dadi, K., Davis, C.P., Luca, A., Delgado, M.R., Demetriou, L., Dennison, J.B., Di, X., Dickie, E.W., Dobryakova, E., Donnat, C.L., Dukart, J., Duncan, N.W., Durnez, J., Eed, A., Eickhoff, S.B., Erhart, A., Fontanesi, L., Fricke, G.M., Fu, S., Galvan, A., Gau, R., Genon, S., Glatard, T., Glerean, E., Goeman, J.J., Golowin, S.A.E., Gonzalez-Garcia, C., Gorgolewski, K.J., Grady, C.L., Green, M.A., et al., 2020. Variability in the analysis of a single neuroimaging dataset by many teams. Nature 582, 84–88.

Bouchard, K.E., Aimone, J.B., Chun, M., Dean, T., Denker, M., Diesmann, M., Donofrio, D.D., Frank, L.M., Kasthuri, N., Koch, C., Ruebel, O., Simon, H.D., Sommer, F.T., Prabhat, 2016. High-performance computing in neuroscience for data–driven discovery, integration, and dissemination. Neuron 92, 628–631.

Bzdok, D., Yeo, B.T.T., 2017. Inference in the age of big data: future perspectives on neuroscience. Neuroimage 155, 549–564.

Chen, A.C., Feng, W., Zhao, H., Yin, Y., Wang, P., 2008. EEG default mode network in the human brain: spectral regional field powers. Neuroimage 41, 561–574.

Cohen, M.X., 2017. Where does eeg come from and what does it mean? Trends Neurosci. 40, 208–218.

Cowley, B.U., Korpela, J., Torniainen, J., 2017. Computational testing for automated preprocessing: a Matlab toolbox to enable large scale electroencephalography data processing. PeerJ Comput. Sci. 3, e108.

Dale, A.M., Liu, A.K., Fischl, B.R., Buckner, R.L., Belliveau, J.W., Lewine, J.D., Halgren, E., 2000. Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. Neuron 26, 55–67.

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134, 9–21.

Dong, L., Gong, D., Valdes-Sosa, P.A., Xia, Y., Luo, C., Xu, P., Yao, D., 2014. Simultaneous EEG-fMRI: trial level spatio-temporal fusion for hierarchically reliable information discovery. Neuroimage 99, 28–41.

Dong, L., Li, F.L., Liu, Q., Wen, X., Lai, Y.X., Xu, P., Yao, D.Z., 2017. MATLAB toolboxes for reference electrode standardization technique (REST) of scalp EEG. Front. Neurosci. 11.

Dong, L., Luo, C., Liu, X.B., Jiang, S.S., Li, F.L., Feng, H.S., Li, J.F., Gong, D.K., Yao, D.Z., 2018. Neuroscience information toolbox: an open source toolbox for EEG-fMRI multimodal fusion analysis. Front. Neuroinform. 12.

Dong, L., Wang, P., Bin, Y., Deng, J., Li, Y., Chen, L., Luo, C., Yao, D., 2015a. Local multimodal serial analysis for fusing EEG-fMRI: a new method to study familial cortical myoclonic tremor and epilepsy. IEEE Trans. Auton. Mental Dev. 7, 1 -1.

Dong, L., Zhang, Y., Zhang, R., Zhang, X., Gong, D., Valdes-Sosa, P.A., Xu, P., Luo, C., Yao, D., 2015b. Characterizing nonlinear relationships in functional imaging data using eigenspace maximal information canonical correlation analysis (emiCCA). Neuroimage 109, 388–401.

Dong, L., Zhao, L., Zhang, Y., Yu, X., Li, F., Li, J., Lai, Y., Liu, T., Yao, D., 2021. Reference electrode standardization interpolation technique (RESIT): a novel interpolation method for scalp EEG. Brain Topogr. 34, 403–414.

Dragoni, N., Giallorenzo, S., Lafuente, A.L., Mazzara, M., Montesi, F., Mustafin, R., Safina, L., 2017. Microservices: Yesterday, Today, and Tomorrow. Present and Ulterior Software Engineering. Springer, pp. 195–216.

Enriquez-Geppert, S., Huster, R.J., Herrmann, C.S., 2017. EEG-neurofeedback as a tool to modulate cognition and behavior: a review tutorial. Front. Hum. Neurosci. 11 (51).

Felter, W., Ferreira, A., Rajamony, R., Rubio, J., 2015. An updated performance comparison of virtual machines and linux containers. In: Proceedings of the IEEE International Symposium on Performance Analysis and Software (ISPASS), pp. 171–172.

Fowler, M., Lewis, J., 2014. Microservices a definition of this new architectural term. MartinFowler. com, 2014, 25: 14-26. URL: http://martinfowler.com/articles/microservices.html, 22.

Frauscher, B., von Ellenrieder, N., Zelmann, R., Dolezalova, I., Minotti, L., Olivier, A., Hall, J., Hoffmann, D., Nguyen, D.K., Kahane, P., Dubeau, F., Gotman, J., 2018. Atlas of the normal intracranial electroencephalogram: neurophysiological awake activity in different cortical areas. Brain 141, 1130–1144.

Friston, K.J., Preller, K.H., Mathys, C., Cagnan, H., Heinzle, J., Razi, A., Zeidman, P., 2019. Dynamic causal modelling revisited. Neuroimage 199, 730–744.

Gabard-Durnam, L.J., Mendez Leal, A.S., Wilkinson, C.L., Levin, A.R., 2018. The harvard automated processing pipeline for electroencephalography (HAPPE): standardized processing software for developmental and high-artifact data. Front. Neurosci. 12, 97.

Glasser, M.F., Smith, S.M., Marcus, D.S., Andersson, J.L., Auerbach, E.J., Behrens, T.E., Coalson, T.S., Harms, M.P., Jenkinson, M., Moeller, S., Robinson, E.C., Sotiropoulos, S.N., Xu, J., Yacoub, E., Ugurbil, K., Van Essen, D.C., 2016. The Human Connectome Project's neuroimaging approach. Nat. Neurosci. 19, 1175–1187.

Goldberger, A.L., Amaral, L.A., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K., Stanley, H.E., 2000. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation 101, E215–E220.

He, B., Gao, S., Yuan, H., Wolpaw, J.R., 2013. Brain-computer interfaces. In: Neural Engineering. Springer, pp. 87–151.

Jobert, M., Wilson, F.J., Roth, T., Ruigt, G.S., Anderer, P., Drinkenburg, W.H., Bes, F.W., Brunovsky, M., Danker-Hopfe, H., Freeman, J., van Gerven, J.M., Gruber, G., Kemp, B., Klosch, G., Ma, J., Penzel, T., Peterson, B.T., Schulz, H., Staner, L., Saletu, B., Svetnik, V., Committee, I.P.-E.G., 2013. Guidelines for the recording and evaluation of pharmaco-sleep studies in man: the international pharmaco-EEG society (IPEG). Neuropsychobiology 67, 127–167.

Koch, C., Jones, A., 2016. Big science, team science, and open science for neuroscience. Neuron 92, 612–616.

Laufs, H., 2012. A personalized history of EEG-fMRI integration. Neuroimage 62, 1056–1067.

Li, F., Wang, J., Liao, Y., Yi, C., Jiang, Y., Si, Y., Peng, W., Yao, D., Zhang, Y., Dong, W., Xu, P., 2019a. Differentiation of Schizophrenia by Combining the Spatial EEG Brain Network Patterns of Rest and Task P300. IEEE Trans. Neural Syst. Rehabil. Eng. 27, 594–602.

Li, F., Yi, C., Song, L., Jiang, Y., Peng, W., Si, Y., Zhang, T., Zhang, R., Yao, D., Zhang, Y., Xu, P., 2019b. Brain Network Reconfiguration During Motor Imagery Revealed by a Large-Scale Network Analysis of Scalp EEG. Brain Topogr. 32, 304–314.

Li, F.L., Liu, T.J., Wang, F., Li, H., Gong, D.K., Zhang, R., Jiang, Y., Tian, Y., Guo, D.Q., Yao, D.Z., Xu, P., 2015. Relationships between the resting-state network and the P3: evidence from a scalp EEG study. Sci. Rep. 5 (1), 1–10.

Lin, Z., Chen, M., Ma, Y., 2010. The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices. arXiv preprint arXiv:1009.5055.

Majumdar, A., Sivagnanam, S., Carnevale, N.T., Yoshimoto, K., Gleeson, P., Quintana, A., Silver, R.A., 2016. Neuroscience gateway-cyberinfrastructure providing supercomputing resources for large scale computational neuroscience research. Front. Neuroinform. 1. doi:10.3389/conf.fninf.2016.20.00008, Conference Abstract: Neuroinformatics.

Makkie, M., Li, X., Quinn, S., Lin, B., Ye, J., Mon, G., Liu, T., 2019. A distributed computing platform for fMRI big data analytics. IEEE Trans. Big Data 5, 109–119.

Malver, L.P., Brokjaer, A., Staahl, C., Graversen, C., Andresen, T., Drewes, A.M., 2014. Electroencephalography and analgesics. Br. J. Clin. Pharmacol. 77, 72–95.

Manjon, J.V., Coupe, P., 2016. volBrain: an online MRI brain volumetry system. Front. Neuroinform. 10, 30.

Marcus, D.S., Olsen, T.R., Ramaratnam, M., Buckner, R.L., 2007. The extensible neuroimaging archive toolkit: an informatics platform for managing, exploring, and sharing neuroimaging data. Neuroinformatics 5, 11–34.

Martinez-Cancino, R., Delorme, A., Truong, D., Artoni, F., Kreutz-Delgado, K., Sivagnanam, S., Yoshimoto, K., Majumdar, A., Makeig, S., 2021. The open EEGLAB portal Interface: high-Performance computing with EEGLAB. Neuroimage 224, 116778.

Merkel, D., 2014. Docker: lightweight linux containers for consistent development and deployment. Linux J. 239, 1–2.

Mognon, A., Jovicich, J., Bruzzone, L., Buiatti, M., 2011. ADJUST: an automatic EEG artifact detector based on the joint use of spatial and temporal features. Psychophysiol. 48, 229–240.

Mullen, T., Kothe, C., Chi, Y.M., Ojeda, A., Kerth, T., Makeig, S., Cauwenberghs, G., Jung, T.P., 2013. Real-time modeling and 3D visualization of source dynamics and connectivity using wearable EEG. In: Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, pp. 2184–2187.

Mullen, T.R., Kothe, C.A., Chi, Y.M., Ojeda, A., Kerth, T., Makeig, S., Jung, T.P., Cauwenberghs, G., 2015. Real-time neuroimaging and cognitive monitoring using wearable dry EEG. IEEE Trans. Biomed. Eng. 62, 2553–2567.

Netto, H.V., Lung, L.C., Correia, M., Luiz, A.F., de Souza, L.M.S., 2017. State machine replication in containers managed by Kubernetes. J. Syst. Arch. 73, 53–59.

Neuro Cloud Consortium, 2016. To the Cloud! A grassroots proposal to accelerate brain science discovery. Neuron 92, 622–627.

Nichols, T.E., Das, S., Eickhoff, S.B., Evans, A.C., Glatard, T., Hanke, M., Kriegeskorte, N., Milham, M.P., Poldrack, R.A., Poline, J.B., Proal, E., Thirion, B., Van Essen, D.C., White, T., Yeo, B.T., 2017. Best practices in data analysis and sharing in neuroimaging using MRI. Nat. Neurosci. 20, 299–303.

Obeid, I., Picone, J., 2016. The temple university hospital EEG data corpus. Front. Neurosci. 10, 196.

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. 2011, 156869.

Pascual-Marqui, R.D., 2002. Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. Methods Find Exp. Clin. Pharmacol. 24 Suppl D, 5–12.

Pedroni, A., Bahreini, A., Langer, N., 2019. Automagic: standardized preprocessing of big EEG data. Neuroimage 200, 460–473.

Pernet, C.R., Appelhoff, S., Gorgolewski, K.J., Flandin, G., Phillips, C., Delorme, A., Oostenveld, R., 2019. EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. Sci. Data 6, 103.

Perrin, F., Pernier, J., Bertrand, O., Echallier, J.F., 1989. Spherical splines for scalp potential and current density mapping. Electroencephalogr. Clin. Neurophysiol. 72, 184–187.

Rubinov, M., Sporns, O., 2010. Complex network measures of brain connectivity: uses and interpretations. Neuroimage 52, 1059–1069.

Schalk, G., McFarland, D.J., Hinterberger, T., Birbaumer, N., Wolpaw, J.R., 2004. BCI2000: a general-purpose brain-computer interface (BCI) system. IEEE Trans. Biomed. Eng. 51, 1034–1043.

Sejnowski, T.J., Churchland, P.S., Movshon, J.A., 2014. Putting big data to good use in neuroscience. Nat. Neurosci. 17, 1440–1441.

Sherif, T., Rioux, P., Rousseau, M.E., Kassis, N., Beck, N., Adalat, R., Das, S., Glatard, T., Evans, A.C., 2014. CBRAIN: a web-based, distributed computing platform for collaborative neuroimaging research. Front. Neuroinform. 8, 54.

Smith, S.M., Nichols, T.E., 2018. Statistical challenges in "Big Data" human neuroimaging. Neuron 97, 263–268.

Tian, Y., Xu, W., Zhang, H., Tam, K.Y., Zhang, H., Yang, L., Li, Z., Pang, Y., 2018. The scalp time-varying networks of N170: reference, latency, and information flow. Front. Neurosci. 12, 250.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15, 273–289.

Van Horn, J.D., Toga, A.W., 2014. Human neuroimaging as a "Big Data" science. Brain Imaging Behav. 8, 323–331.

Winkler, I., Haufe, S., Tangermann, M., 2011. Automatic classification of artifactual ICA–components for artifact removal in EEG signals. Behav. Brain Funct. 7, 30.

Xu, P., Xiong, X., Xue, Q., Li, P., Zhang, R., Wang, Z., Valdes-Sosa, P.A., Wang, Y., Yao, D., 2014. Differentiating between psychogenic nonepileptic seizures and epilepsy based on common spatial pattern of weighted EEG resting networks. IEEE Trans. Biomed. Eng. 61, 1747–1755.

Yamada, T., Meng, E., 2012. Practical Guide For Clinical Neurophysiologic testing: EEG. Lippincott Williams & Wilkins.

Yao, D., 2001. A method to standardize a reference of scalp EEG recordings to a point at infinity. Physiol. Meas. 22, 693–711.

Yao, D., Wang, L., Oostenveld, R., Nielsen, K.D., Arendt-Nielsen, L., Chen, A.C., 2005. A comparative study of different references for EEG spectral mapping: the issue of the neutral reference and the use of the infinity reference. Physiol. Meas. 26, 173–184.

Yao, D.Z., Qin, Y., Hu, S., Dong, L., Vega, M.L.B., Sosa, P.A.V., 2019. Which reference should we use for EEG and ERP practice? Brain Topogr. 32, 530–549.

Zhang, R., Li, X., Wang, Y., Liu, B., Shi, L., Chen, M., Zhang, L., Hu, Y., 2019. Using brain network features to increase the classification accuracy of MI-BCI inefficiency subject. IEEE Access 7, 74490–74499.

Zhao, X., Liu, C., Xu, Z., Zhang, L., Zhang, R., 2020. SSVEP Stimulus Layout Effect on Accuracy of Brain-computer interfaces in Augmented Reality Glasses. IEEE Access.