Medical Data Analytics: Human intracerebral EEG Platform (HIP) and Medical Informatics Platform (MIP) (SC5) - status at M21 (D4.9 - SGA3)

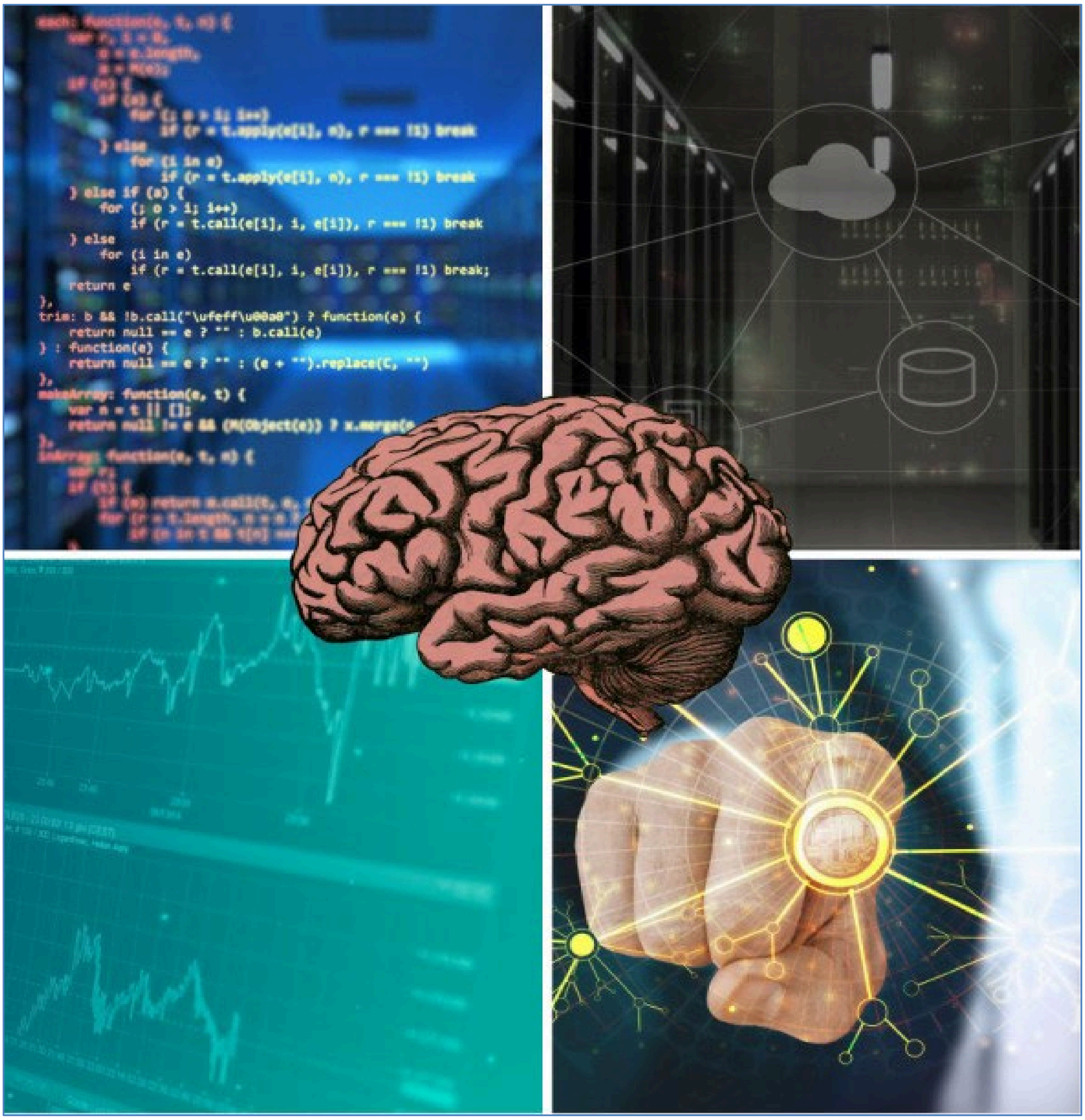

Figure 1: EBRAINS Medical Data Analytics[1]

Abstract representation of the Medical Data Analytics Service, representing the Medical Informatics Platform (MIP) and the Human intracerebral EEG Platform (HIP), Services focussed on the area of Brain Health.

[1] Image sources: brain: https://pixabay.com/fr/illustrations/cerveau-anatomie-humain-science-512758/; code top left: https://pixabay.com/photos/source-code-software-computer-4280758/; graph bottom left: https://pixabay.com/photos/business-chart-graph-graphic-5475664/; cloud top right : https://pixabay.com/photos/network-server-system-2402637/; pointing bottom right : https://pixabay.com/photos/web-network-technology-developer-3963945/

Project Number:	945539	Project Title:	HBP SGA3

Document Title:	Medical Data Analytics: Human intracerebral EEG Platform (HIP) and Medical Informatics Platform (MIP) (SC5) - status at M21
Document Filename:	D4.9 (D40) SGA3 M21 ACCEPTED 220520.docx
Deliverable Number:	SGA3 D4.9 (D40)
Deliverable Type:	Other
Dissemination Level:	PU = Public
Planned Delivery Date:	SGA3 M21 / 31 Dec 2021
Actual Delivery Date:	SGA3 M22 / 10 Jan 2022; accepted 20 May 2022
Author(s):	Birgit SCHAFFHAUSER, Emrah KAVUN, CHUV (P27); Olivier DAVID, AMU (P78)/UGA (P125); Marmaduke WOODMAN AMU (P78); Giorgos PAPANIKOS, Evita MAILI, ATHENA (P133)
Compiled by:	Birgit SCHAFFHAUSER, CHUV (P27)
Contributor(s):	Olivier DAVID, Anthony BOYER, Manik BHATTACHARJEE, AMU (P78)/UGA (P125), contributed to Sections 3 and 4 Philippe RYVLIN, Emrah KAVUN, Erika BORCEL, Manuel SPUHLER, Steve REIS, Cedric ROCHAT, Laith ABU-NAWWAS, CHUV (P27), contributed to Sections 1, 2, 3, 4 Carolina CIUMAS, Nathalie CASATI, CHUV (P27), contributed to Section 3 Viktor JIRSA, Marmaduke WOODMAN, Jan FOUSEK, AMU (P78), contributed to Section 3 Jean-Philippe LACHAUX, Florian SIPP, Benjamin BONTEMPS, UCBL (P108), contributed to Section 3 Yannis IOANNIDIS, Giorgos PAPANIKOS, Evita MAILI, ATHENA (P133), contributed to Sections 2 and 4 Vasilis VASSALOS, Iosif SPARTALIS, AUEB (P4), contributed to Sections 2 and 4
WP QC Review:	Roman VOLCHENKOV, UIO (P81)
WP Leader / Deputy Leader Sign Off:	Jan BJAALIE, UIO (P81)
T7.4 QC Review:	Martin TELEFONT, Annemieke MICHELS, EBRAINS (P1)
Description in GA:	Title: Medical Data Analytics: Human intracerebral EEG Platform (HIP) and Medical Informatics Platform (MIP) (SC5) - status at M21 Description: First release of HIP platform with all services needed for end-users to upload, store and curate iEEG data.
Abstract:	The EBRAINS Medical Data Analytics Service offers approaches to make data sharing and data federation in the fields of brain health, clinical neuroscience and medicine in general, more accessible to the clinical research community. It provides secure solutions for federated and centralised data processing, and contributes to the digital transformation and globalisation of research in medicine. Two EBRAINS platforms, the Medical Informatics Platform (MIP) and the Human Intracerebral EEG Platform (HIP), attend to address real needs and challenges in

	harmonising, sharing, federating, and analysing complex brain-health data. This status report provides an overview of achievements and ongoing developments for both platforms.
Keywords:	Data federation, federated algorithms, data interoperability, code-visits-data, virtual appliance, data model, data privacy, human iEEG data, 3-tier architecture, In-Browser-Apps, data harmonisation, brain health, clinical neuroscience, medical data
Target Users/Readers:	clinicians, computational neuroscience community, computer scientists, HBP Consortium members, general public, neuroinformaticians, neuroscientific community, neuroscientists, platform users, researchers

Table of Contents

1. Introduction - Medical Data Analytics (SC5) 5
2. The Medical Informatics Platform - MIP 6
2.1 Medical Informatics Platform maintenance and upgrade 6
2.1.1 Evolution of the MIP technology stack 6
2.1.2 Usability and accessibility enhancements of the MIP User interface 7
2.1.3 MIP scalability enhancements 8
2.1.4 MIP deployment 8
2.1.5 EBRAINS embedded MIP federations 9
2.1.6 MIP Releases – latest release MIP6.4 10
2.1.7 MIP Data Governance 11
2.2 MIP-embedded digital twin architecture – Prototype design 14
2.3 MIP Exploitation 16
2.3.1 The PUBLIC MIP 16
2.3.2 The MIP Hands-On Training Federation 16
2.4 Ongoing developments - current work on the MIP 16
2.4.1 New MIP Analysis Engine 16
2.4.2 MIP New Federations and Collaborations 17
2.4.3 New MIP release - MIP6.5 18
3. The Human intracerebral EEG platform - HIP 21
3.1 The Human intracerebral EEG data curation and management 21
3.1.1 The HIP Network 21
3.1.2 The HIP Legal and Ethical Framework 22
3.1.3 The HIP 3-tier architecture 23
3.1.4 HIP Storage and Compute 26
3.1.5 The HIP standardised procedures and related tutorials 26
3.2 The Human intracerebral EEG data and analytical tools integration 27
3.2.1 The HIP Library – In-Browser Apps 27
3.2.2 The integration of BIDS-iEEG manager software tools 28
3.2.3 Integration of visualisation tools 29
3.3 The Human intracerebral EEG data services: integration with other EBRAINS Services 31
3.3.1 The HIP structural and functional processing workflows 31
3.3.2 TVB workflows and pipelines 32
3.3.3 HIP Adaptors 32
3.3.4 The HIP Access, Use and Accreditation Modalities 32
3.3.5 Integration of F-Tract in the Multilevel Human Brain Atlas 33
3.3.6 Integration of new HIP apps and pipelines 33

3.4 The HIP Prototype – platform status in December 2021 34
3.4.1 The HIP repositories and Back end 34
3.4.2 The HIP Front end and Gateway 34
4. Looking Forward 35
4.1 MIP future developments 35
4.2 HIP future developments 35
5. ANNEX 36

Table of Tables

Table 1: European registries interested in participating in the Stroke federation 17
Table 2: Centres participating in the NeuroCohort pilot federation 18
Table 3: The HIP Library – In-Browser Apps 27
Table 4: Documentation and other references 36

Table of Figures

Figure 1: EBRAINS Medical Data Analytics 1
Figure 2: The MIP6.4 User Interface 7
Figure 3: The MIP Network Topology 9
Figure 4: Architecture of a hybrid Federation in the MIP 10
Figure 5: The MIP Data Governance flow 13
Figure 6: User interface of the MIP Data Quality Tool 14
Figure 7: The MIP data architecture and flow 15
Figure 8: MIP high-level architecture with addition of the SMPC subsystem. 17
Figure 9: Modular MIP architecture with Gateway approach 20
Figure 10: Results of the survey conducted with leading SEEG centres worldwide. 21
Figure 11: The HIP system architecture – a high-level overview 25
Figure 13: A HIP session, running FSL, displaying other available apps in the sidebar 29
Figure 14: Anywave and Brainstorm in-browser apps opened in a single session on the HIP 30
Figure 15: Data visualisation on the HIP: Use of the HiBoP app for 3D visualisation of evoked potentials 30
Figure 16: TVB-HIP in browser app and its relation to other components and data flows. 31

1. Introduction - Medical Data Analytics (SC5)

The EBRAINS Medical Data Analytics Service addresses growing needs in harmonising, sharing, federating, and analysing complex brain-health data. It provides innovative solutions to security and IT challenges, promotes collaboration, and works to accelerate progress in brain health, clinical neuroscience and medicine in general. The Medical Data Analytics Services (SC5) make data sharing and data federation more accessible, and contribute to the digital transformation of research and medicine.

The services offer distinct solutions for federated and centralised data processing, with technologies and algorithms accessible online, in a browser, and with minimal user and IT requirements. Two unique EBRAINS platforms, the Medical Informatics Platform (MIP) and the Human Intracerebral EEG Platform (HIP), co-designed with HBP developers, researchers and users, serve to address real needs and challenges of the clinical neuroscience community. Their adoption in clinical settings, with real use case data, will allow paving the way to future data standards, federation methodologies, and collaborative solutions in digital medicine.

Building on previous efforts by the Human Brain Project (HBP) in SGA1 and SGA2, the concept of data federation and federated data analytics, as implemented in the MIP, offers a powerful solution to the dualism of data sharing versus privacy, augmenting research potential and data value of clinical information, that cannot otherwise be transferred or centralised, e.g. based on institutional or national regulations. Notably, data privacy and ethics policies impose stringent restrictions on sharing medical data, with policies that evolve with time, and often remain specific to a country, territory or jurisdiction. The MIP, with its defined governance framework and the implementation of standardised harmonisation rules, is a research tool optimised and secured for federated analysis. The solution enables hospitals and other clinical research institutions to leverage, explore and analyse medical data, and offers new potential to the challenge and large volume of unexploited clinical data. The solution can help mitigating difficulties inherent to limited sample sizes in certain medical areas for example. The technology has been designed to be minimally invasive in an IT, ethics and legal context. By deploying this technology as virtual appliances, it has become one of the easiest solutions to onboard in hospitals and clinical centres.

The Human Intracerebral EEG Platform (HIP) represents a unique, open-source European platform dedicated to the centralisation of Human iEEG data, which are only recorded in patients with drug-resistant focal epilepsy or after deep brain stimulation, available in a small number of clinical centres. The platform encompasses a collection of tools and services, designed to foster collaborative processing and analysis of data, while providing guidance and scientific pipelines to promote interoperability, comprising data curation and harmonisation.

The development of the HIP is a promising, innovative project that started in SGA3 aiming at centralising and facilitating collaboration across heterogeneous and fragmented iEEG clinical data. This service will offer a fine-grained ecosystem for the collection, storage, curation, sharing and analysis of iEEG data, a first of its kind. It will provide accredited clinical researchers unprecedented access to ethics and GDPR-compliant resources, expected to reach an aggregate of more than 70 clinical centres worldwide, with data collated in curated, multi-scale neurophysiological iEEG databases. The embedded library of supporting domain-specific tools (in-browser-apps) and workflows, and the integration of various HIP functionalities with other EBRAINS services, e.g. the EBRAINS Knowledge Graph (Data and Knowledge), the Multilevel Human Brain Atlas (Atlases), the Virtual Brain (Simulation) with Showcase 2 (Improving epilepsy surgery with the Virtual Big Brain) or the EBRAINS HPC centres, will provide users with an innovative and exceptional experience of a desktop-like environment, minimising the complexity of iEEG and neuroimaging data curation, processing and analysis.

Sections 2 and 3 of this document describe the current functionalities and ongoing developments of the MIP and HIP platforms respectively. Section 4 covers future developments envisaged for the EBRAINS Medical Data Analytics Service and Section 5 (Annex) provides a summary of relevant links to documentation, tools and other material around the MIP and the HIP, as well as the MIP Legal Agreements currently in use.

2. The Medical Informatics Platform - MIP

2.1 Medical Informatics Platform maintenance and upgrade

The MIP links brain science, clinical research and patient care and promotes collaborative research. In a unique and accessible space, it provides the tools to improve knowledge, diagnosis, early prediction and innovative treatment of brain diseases. Being a federated data processing and analysis system, it is best suited for the exploration and modelling of de-centralised anonymised and harmonised medical data, extracted from pre-processed neuroimaging, neurophysiological and medical records and research cohort datasets.

The main principle of the MIP is "Code visits Data". The Central MIP engine hosted on the EBRAINS RI, enables privacy-aware federated web-based queries, which will execute machine learning algorithms in each of the selected nodes/hospitals, by breaking the algorithms in global (central node) and local (remote node) steps. Accredited end-users query federated datasets on remote nodes and retrieve aggregated findings. Importantly, the system prohibits the investigation of datasets at the individual level. Databases cannot be directly accessed, copied, uploaded or downloaded; clinical raw data neither leave their original site of storage, nor are those data uploaded to the MIP.

The focus for the MIP in SGA3 is to strengthen and expand the hospital network, the maintenance and servicing of the current MIP instances and to work towards a sustainable roadmap that reaches beyond the end of SGA3. Further, the MIP will provide additional privacy guarantees by exploring privacy methods, such as Secure Multi Party Computation (SMPC) and Differential Privacy (DP). The sections below reflect the current status and developments to achieve the set goals.

2.1.1 Evolution of the MIP technology stack

The implementation of a semi-automated procedure for MIP installations was a major update and improvement of the MIP compared to the solution available at the end of SGA2. Vagrant[2], a tool for building and managing virtual machine environments with a focus on automation, was implemented. Various system execution aspects were scripted and automated, (e.g. dependencies management, component download (system package, github, dockerhub, etc.), system permissions passing, configuration parameters for machine-to-machine interaction or real-time monitoring of federated nodes, re-start of nodes), which led to a drastic reduction of required manual interventions.

The overall MIP security system consists of a secure virtual private network (VPN) layer using encryption over a public network, running on Ubuntu 20.04 and the Debian GNU[3]/Linux open-source operating system. An improved solution with configuration enhancements for the reverse-proxy server, embedded in the portal front-end container, and the EBRAINS Identity and Access management, (i.e. EBRAINS accreditation and Keycloak authentication, delineating, implementing and controlling the access rights of each accredited end-user of the federation), provides programmatic secure access for each user.

Improvements based on the implementation of a VPN (implemented at the transition from SGA2 to SGA3, with further refinements in SGA3):

- Strong authentication for hospitals
- Over-the-top encryption of the data channels, encapsulating and securing all exchanges
- All connections from hospitals are outbound, no port-forwarding required

[2] Vagrant: https://www.vagrantup.com/
[3] DEBIAN GNU https://www.debian.org/

- Smart with ports and firewall friendly: the VPN can use TCP or UDP
- Protocols, and negotiate among various standard ports (i.e. 443 for HTTPS browsing)

Central Proxy and Firewall:

No federated MIP is exposed directly to the public Internet. The central nodes provide a firewall between client connections and a proxy, where needed. The embedded MIP proxy redirects the user to the main URL upon any error. An analysis is dispatched, distributed and coordinated by the central nodes, which also render the aggregate results, never compromising the data of an individual MIP. All analyses and outputs are systematically recorded and associated to the user's credentials, providing full traceability of data access and data analyses.

Secure storage:

- Virtual isolation using different MIP volumes to maintain the security of data and keep audit trails of back-end access.

2.1.2 Usability and accessibility enhancements of the MIP User interface

User Interface and User Experience are crucial for defining whether a solution is taken up by end-users. Services often focus on the needs of advanced users only. This is where the MIP with a front end as seen in Figure 2, especially developed to serve the clinical community, is a major asset. The design of the UI allows to adopt advanced statistical analytics based on the integrated statistical methods and predictive machine-learning algorithms for data exploration and modelling, diagnosis and research in clinics.

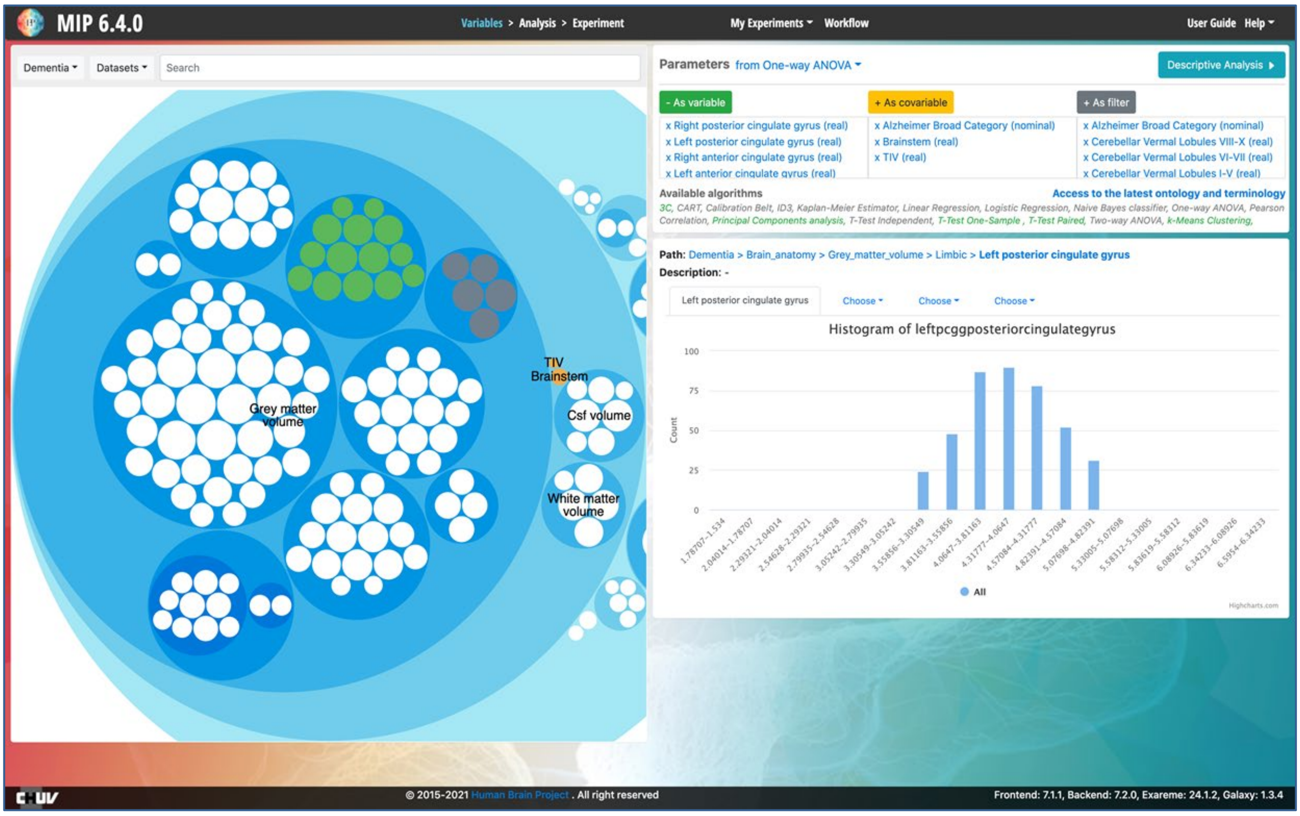

Figure 2: The MIP6.4 User Interface

Notably, the UI setup does not require users to have expert knowledge in programming or design of algorithms; sharing of experiments and workflows between users is enabled; authorisation and security is applied uniformly and vertically across the enabling and presentation layers. The currently offered features and easily accessible design have attracted community interest and have led to new collaborations. An example is the IMI SOPHIA, where the MIP became a partner and where the front-end module is now investigated and further developed for its use in the field of obesity. See also Section 2.1.6.

2.1.3 MIP scalability enhancements

The expansion of the MIP network, one major goal in SGA3, goes along with increased data volumes, higher request rates, demand for higher system capacity to run concurrent analysis and the ability to keep the perceived response time by the end-user within acceptable margins. Measures have been taken to be ready for a growing user base and to improve the MIP's capability to adapt to an increased workload.

The entire MIP infrastructure is now executed on the EBRAINS Research Infrastructure with a secure VPN to provide common, over-the-top and firewall-friendly network topologies for data transport, remote access, connectivity and infrastructure interoperability across all modules. The MIP operations and technical aspects of the deployment process were automated to a large extent. This was achieved mainly by developing tools to

- reduce the human resources needed to scale-up federations
- assist in remote management of servers of the different hospitals
- enhance installation, upgrade and configuration procedures of the MIP
- automate generation of virtual machines (VMs) that are provided for hospitals
- authorisation configuration bootstrapping automations

Technical documentation: MIP Deployment[4] ; MIP EXAREME[5]

Current work on the development and integration of an API-coupled middleware (see Section 2.4.3.2), which allows the abstraction of interfaces and accepts inputs and outputs of analysis engines based on plugins or dedicated components, will advance the MIP to a more modular architecture. This will enhance the functional scalability of the MIP by easing the addition of new features without disrupting the existing system. Additionally, asynchronous communication and message / task driven approaches incorporated in the system, allow further horizontal scaling while extending the overall system throughput, utilising established Advanced Message Queuing Protocols (AMQP).

The topological structure of the MIP network to date entails around 90 VMs on EBRAINS (i.e. including the federations, QA and development environments, the PUBLIC MIP and the new Hands-on federation for training sessions), and 30 remote nodes installed across Europe. Each federation consists of a “MIP Central” VM with the analysis engine EXAREME and a “Pusher” VM embedded in the EBRAINS RI, which are connected with the federated nodes either hosted by the federated institutions or embedded as a federated VM on EBRAINS, which hold the anonymised datasets and an instance of the analysis engine (see Figure 3).

2.1.4 MIP deployment

The MIP network entails 34 hospitals and institutions in 16 countries across Europe. Twelve institutions signed legal agreements and are awaiting installation, detailed discussions with ten further centres for MIP installation are ongoing, which brings us very close to our target of 60 hospitals with an allocated MIP by the end of SGA3. Installation of a MIP in new hospitals is pending the release of the last MIP version to offer the best possible service and end-user experience. Federations on four brain disorders (Dementia, TBI (traumatic brain injury), Epilepsy, Mental Health) were previously established; new federated nodes can be added to these federations at any time.

[4] MIP Deployment https://github.com/HBPMedical/mip-deployment/blob/master/doc/Readme.md
[5] MIP EXAREME https://github.com/madgik/exareme/tree/24.1.2/Federated-Deploymentbi

The MIP is a data-agnostic platform, this means that interested researchers can start new domain-specific federations in any scientific field, in– and outside of the area of Brain Health.

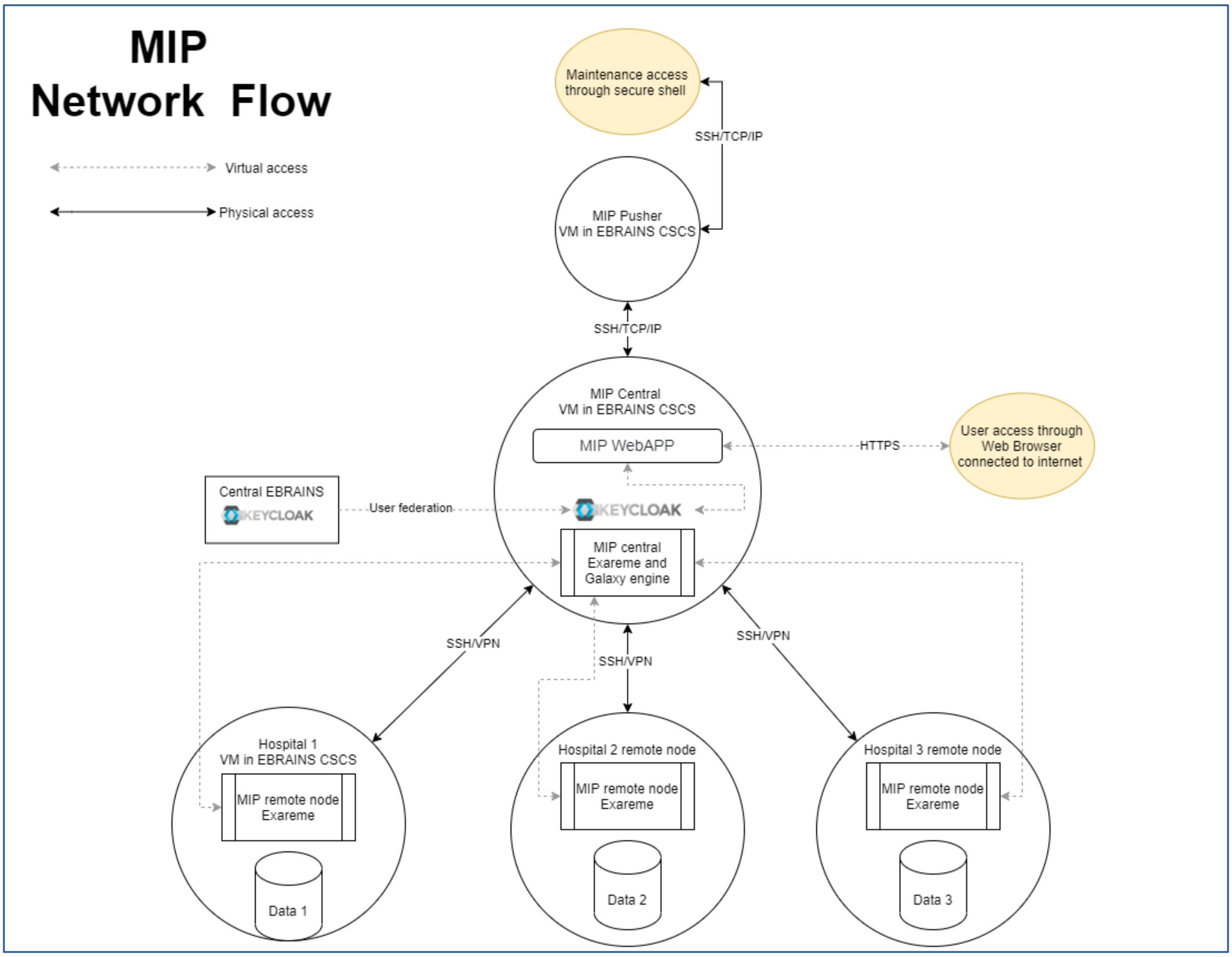

Figure 3: The MIP Network Topology

The deployment process has been streamlined and adjusted to the new deployment options available for the MIP, i.e. a remote node installation within the institution's infrastructure, or alternatively and new in SGA3, a virtual machine (VM) hosted on the EBRAINS RI (see Figure 4). The implementation of a short technical user survey to collect relevant technical information at the start of interactions with new centres improved the effectiveness of the technical user support and reduced the MIP installation time.

MIP Agreements: Updated or new key documents, adapted to the current configuration of the federated MIP, are available: These are *The Medical Informatics Platform Service Agreement and Software Licenses* (mandatory for any MIP deployment), *The Medical Informatics Platform Data Sharing and Processing Agreement* (used for remote instances), and *The Medical Informatics Platform Data Transfer Agreement* (to be used by centres that opt for an EBRAINS virtual machine).

Additionally, a Letter of Understanding, to be used by the participants of a pilot study, was put in place, which outlines the implementation of the use case on testing the MIP in a federation of synthetic data in the field of dementia. A detailed schema for the use of the various agreements is outlined in Figure 5.

2.1.5 EBRAINS embedded MIP federations

As mentioned above, new in SGA3 is the offer of EBRAINS embedded federations, or EBRAINS embedded MIP instances (VMs) as an alternative option to the current model of MIP federation, where MIP federated nodes are installed directly in hospitals. The possibility to use a "private" and secured virtual machine (VM) pre-installed within EBRAINS, provides a welcome solution for users who experience difficulties accessing or providing the required IT infrastructure or IT support within their institutions. Each hospital choosing this option benefits from its own EBRAINS VM with dedicated restricted access for their accredited users only. The three possible models for federations are now a) several remote nodes being federated; b) several VMs on EBRAINS being federated (EBRAINS-embedded federation) and c) a hybrid federation (see Figure 4), where both, remote and hosted nodes are federated within the same federation.

The model of hybrid federations will most likely be the most popular setup of federations, providing partners with the highest degree of freedom, to choose the most suitable solutions. This model is currently taken up by the Epilepsy Federation (see Section 2.4.2).

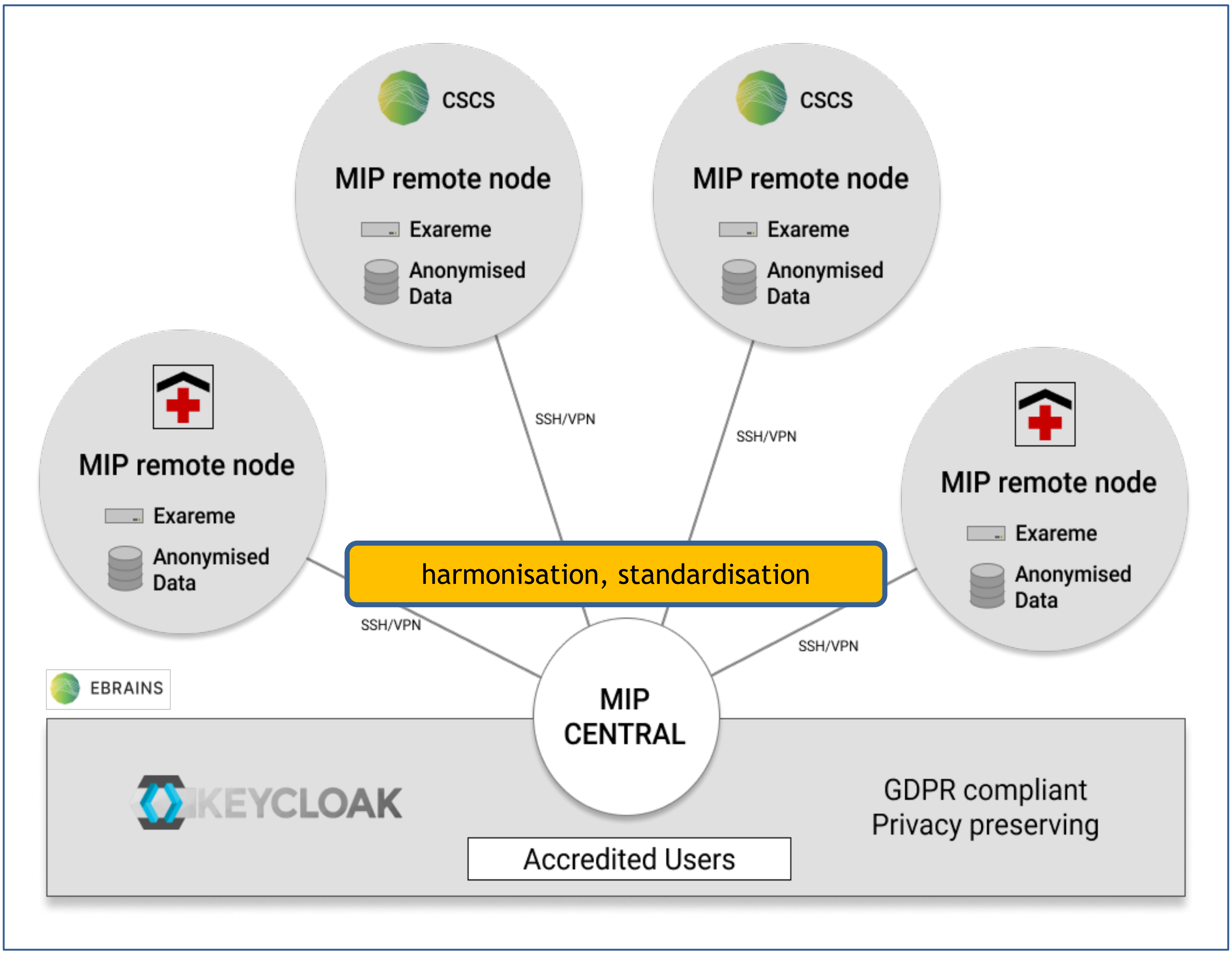

Figure 4: Architecture of a hybrid Federation in the MIP

2.1.6 MIP Releases – latest release MIP6.4

Since the end of SGA2, when the MIP reached version MIP6.0, successive updates, which aimed at improvements in stability, accessibility, usability, the available analyses options, etc., were released. Now arriving at MIP6.4, we achieved to develop an easy to install, user friendly and stable release of the MIP.

A user survey revealed a demand of additional analytics and visualisation capabilities, both for expert users, but also for standard users, who wish to benefit from a larger selection of pre-defined algorithms. Based on this feedback, further enhancements were introduced:

New algorithms:

- One-Way ANOVA with Tukey multiple pairwise comparisons.

Enhanced visualisations and performance:

- Naive Bayes classifier with confusion matrix; CART visualisation showing sample size; Logistic Regression: Confusion matrix output, positive/negative outcome deceleration.

UX enhancements:

- Online user guide updated; experiment module, share, search, delete past experiments, pagination implemented; visual component responsiveness; improvement of the layout and labelling of the "Bubble overview" of variables; integration with the new Back-end API; Web server moved from nginx to Caddy.

Data Management:

- Standardising categorical variables from {multi/poly/bi} to {nominal}; data ingestion pipeline stability enhancements; Quality Control Tool features re-design and user interface enhancements.

Technical documentation: MIP Federated Algorithms[6]

2.1.7 MIP Data Governance

One challenging aspect in working across institutional or even national boundaries, are limitations regarding transparent and FAIR operational frameworks and governance models for the exchange and secondary use of health data. As described before, the MIP offers an easily adoptable solution for trans-national data federation, exploration and analysis, with a legal framework in place that allows institutions to collaborate across boarders, in any field of interest, on any kind of extracted data.

Other challenging aspects are data collection and sharing, often due to the lack of data interoperability, especially in case of electronic health records (EHRs), and thus time-consuming data curation processes. Substantial efforts to ease these aspects were made in SGA3, leading to a semi-automated process for data pre-processing, harmonisation and standardisation in the MIP. Clinical data heterogeneity often requires substantial efforts in standardising and harmonising datasets between institutions (e.g. due to the use of a variety of tools, scales or protocols for measuring the same variables).

Figure 5 depicts how data governance and data flow in the MIP are organised and how legal framework and data management are interlinked.

The respective role of federated institutions versus EBRAINS MIP central nodes, in governing each node's dataset, has been thoroughly discussed in the past. We have always militated for the data holder of each federated institution to remain the sole data controller of its federated dataset, in charge of 1) its compliance with regulatory and ethics requirements, 2) its curation, harmonisation to the federated data model and quality control, and 3) authorisation of other parties to query their dataset through the federation. EBRAINS and the MIP provide the tools and guidance, including dataset-specific Keycloak authentication and accreditation, to enable data controllers assume all their responsibilities. A recent landmark study on federated learning in medicine used this approach: "Governance of data for federated learning is maintained locally" (Dayan et al., Nat Med 2021[7]).

2.1.7.1 MIP Data Management and Quality control

With the MIP, we offer several user-support tools, to enable users to tackle the above-mentioned issues, and to simplify the process of data preparation:

[6] MIP Federated Algorithms https://github.com/madgik/exareme/tree/22.2.3/Exareme-Docker/src/mip-algorithms

[7] Dayan et al., Federated learning for predicting clinical outcomes in patients with COVID-19. Nat Med 2021: https://doi.org/10.1038/s41591-021-01506-3

- The MIP Data Catalogue[8]: A web-based application used to explore and manage MIP Data Models and CDEs pre-defined by the community (in a federation) or the use case group. All the existing Medical Conditions, Federations and Hospitals and their associated Common Data Elements (CDEs) are available within the MIP Data Catalogue, which is the unique source of reference for the MIP and the MIP-DQC Tool.
    - The MRI parallel neuromorphometrics pipeline: A python wrapper to run the Statistical Parametric Mapping SPM12 pipeline in parallel with Matlab over multiple CPU cores.
- The MIP Data Quality Tool (MIP-DQC Tool): Stand-alone software that provides hospital personnel with an easy way to explore, validate and transform their data set based on the CDEs, before uploading it to the MIP. It provides a defined pipeline to extract, standardise, map and harmonise data according to a defined data model. The MIP-DQC Tool has both, a Command Line Interface (CLI) and a Graphical User Interface (GUI). The MIP-DQC Tool GUI version has the following functionalities:
    - Validating the hospital tabular (csv) data, producing a validation report with overall statistics about the data
    - Data cleaning capability based on the received validation of data
    - Inference of a dataset's schema and producing a schema file in a Frictionless json or in the MIP Data Catalogue's excel format
    - Designing and performing schema mapping of an incoming hospital dataset to a certain Common Data Element (CDE) schema
    - Producing DICOM MRI validation and a statistical report based on the meta-data headers
    - The tabular (csv) data validation functionality has the option of downloading CDE metadata directly from the MIP Data Catalogue's API
    - The schema mapping is performed by the MIPMap engine packaged in a Docker container, which runs in the background

Currently, the most time-consuming aspect is the agreement between partners of a federation on a Common Data Element (CDE), i.e. definition of used variables, scales and measurements, etc.). The Data Quality Tool (see Figure 6) supports all the steps afterwards. This process, despite being semi-automated, is strongly guided by the MIP Data Manager, who is central for coordinating these collaborative efforts.

In the federated MIP, all datasets within a specific federation will follow a standardised metadata schema, defined as the CDE or Data model. The CDE file contains the list of distinct variables and is managed through the MIP Data Catalogue (versioned data archive). Starting with the extraction of the CDE from each dataset (in csv format), the merging of all exports in one standardised variable dictionary, is followed by data harmonisation, an automatic integration step to build the data model.

Major efforts to optimise the functionalities of the MIP data management and quality control to enable independence of end-users in data curation, were released in MIP6.4 (see Section 2.1.6). The resulting improved Data Quality Tool supports users in processing, integrating, organising and semantically enriching data originating from heterogeneous sources. It guarantees the import of high-quality data that comply with hospitals' and EU data privacy rules and GDPR in the MIP.

[8] MIP Data Catalogue: https://datacatalogue.hbpmip.link/pathologies

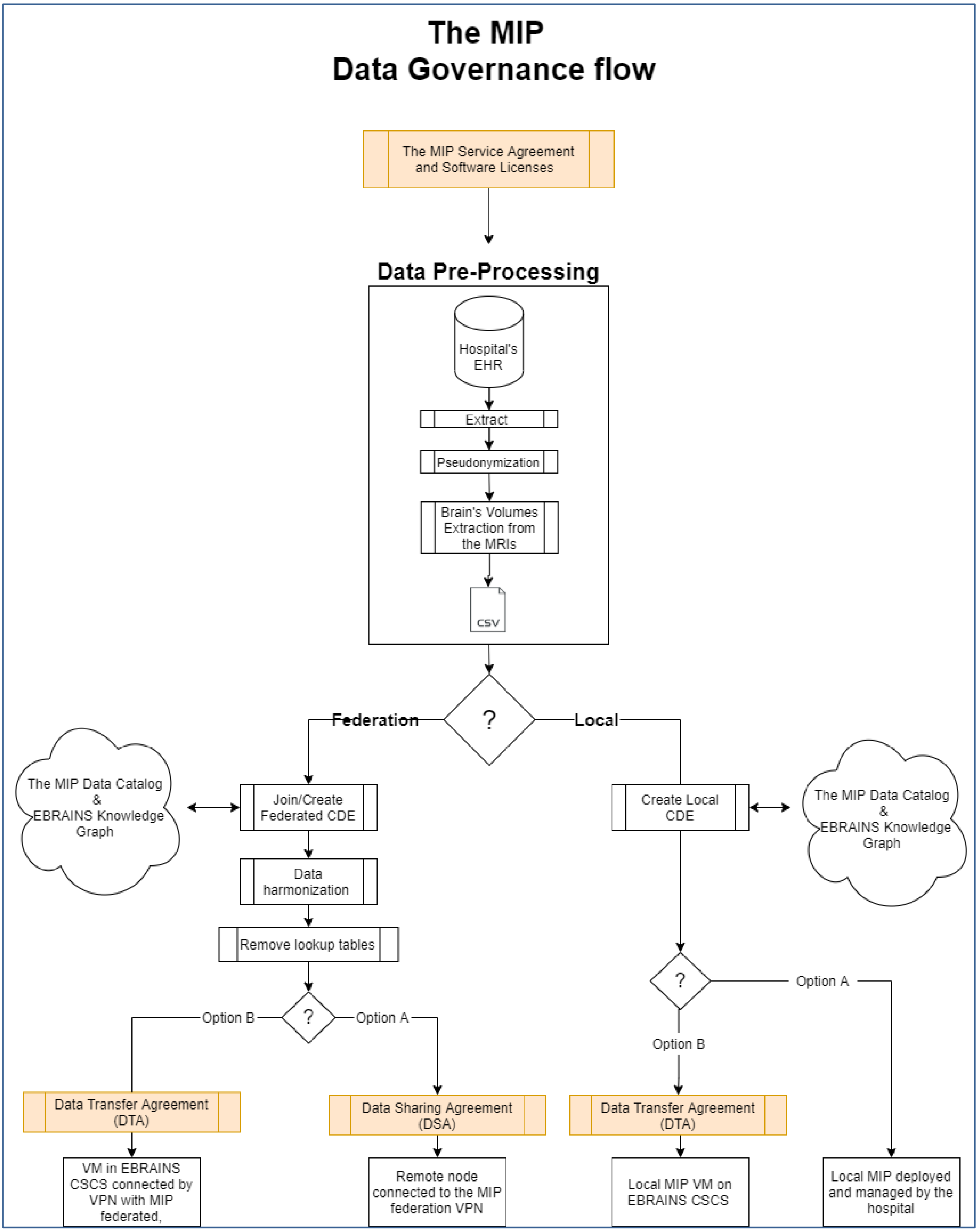

Figure 5: The MIP Data Governance flow

The MIP Data Quality Tool[9] allows to:

- Check the quality of the institute's tabular data against the Common Data Elements created in the MIP Data Catalogue[10].
- Extract the dataset schema to be used as a reference for building the Common Data Elements.

[9] MIP Data Quality Tool https://github.com/HBPMedical/DataQualityControlTool
[10] MIP Data Catalogue https://datacatalogue.hbpmip.link/pathologies

- Design and perform schema mapping using the MIPMap[11] engine.
- Use an optional feature: the DICOM MRI metadata validation.

The MIP Data Catalogue:

- Is the unique source of reference for the Common Data Elements (CDEs) for the various pathologies in the MIP.
- Provides a presentation of the medical conditions and the corresponding Common Data Elements (CDEs) in the MIP for both federated and local use cases.
- Generates metadata files in JSON format using the xlsx extract according to the MIP specifications.

The MIPMap

- Is the engine that transforms hospital tabular data by executing the mappings created by the MIP Data Quality Tool.

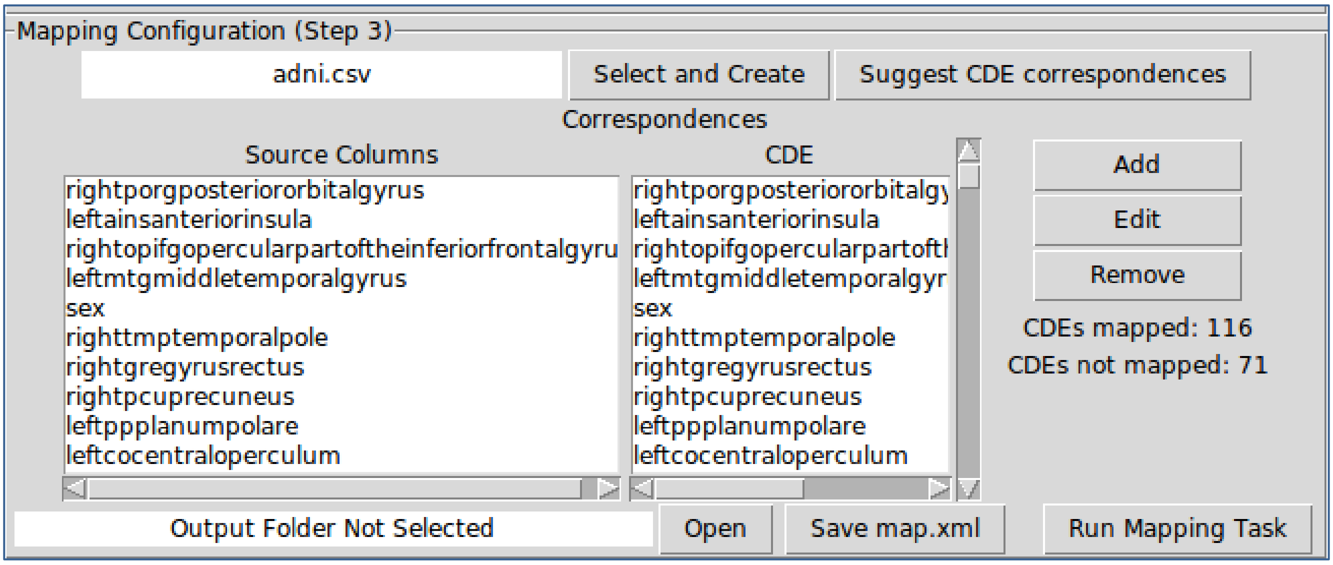

Figure 6: User interface of the MIP Data Quality Tool

2.2 MIP-embedded digital twin architecture – Prototype design

The MIP architecture is based on the premise that strong privacy guarantees should be offered for hospitals to participate in a MIP federation. Mapping the road towards a MIP-embedded digital twin architecture that would enable early diagnosis of rare diseases, we have identified the extension points in the MIP's current architectural approach, along with the changes that this may require to support the use case and shift the focus of the current design to early identification of rare diseases. This architecture modification is challenging, and several considerations should be taken into account, including Authorisation, Data & Metadata, Analysis requirements, User Interface and Visualisation, Access, and Privacy. The above considerations are described in detail in Deliverable D4.4 (D35) *Proposal of MIP architecture for rare disease diagnostic*.

[11] MIPMap https://github.com/HBPMedical/MIPMap

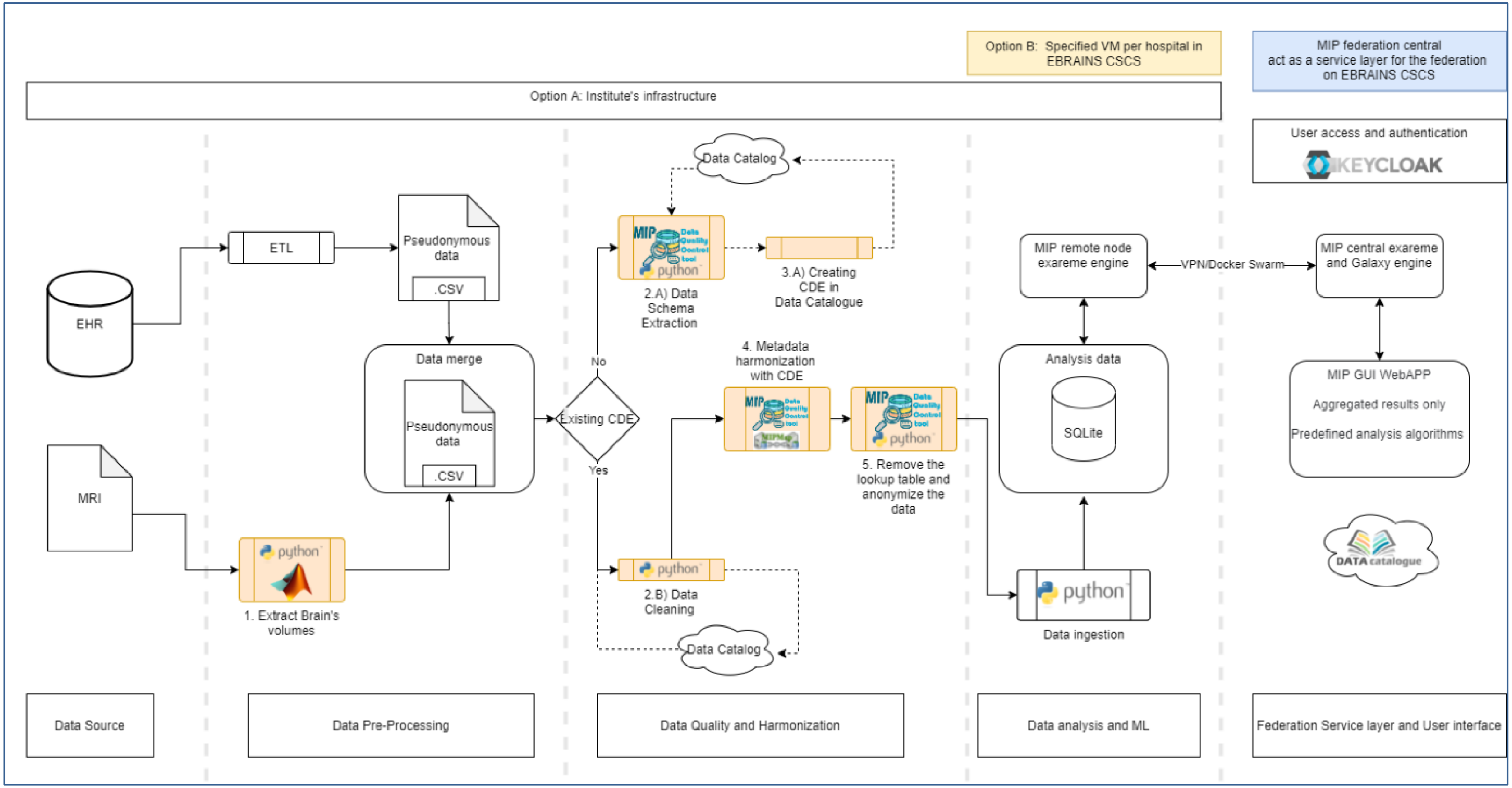

Figure 7: The MIP data architecture and flow

EHR - electronic health record, MRI - magnetic resonance imaging, ETL - data integration (extract, transform, load), CDE - common data elements, ML - machine learning, GUI - graphical user interface, VM - virtual machine.

Data pre-processing: extract data from EHR records and produce pseudonymised data in .csv format; optional Step1: extract brain volumes from MRI images and merge with data extracted from EHR records; Data Quality and Harmonisation: Prepare CDE: if CDE exists - Steps 2B, 4 and 5 are followed; if CDE needs to be prepared, first Steps 2A and 3A need to be performed, followed by Steps 2B, 4 and 5. Data Analysis and ML: anonymised dataset is uploaded either to the remote node or the dedicated VM on EBRAINS CSCS. Data Analysis can be performed via the Federation Service Layer and User Interface: use of predefined federated algorithms, aggregated results will be retrieved via the GUI.

2.3 MIP Exploitation

The MIP is a service, accessible to accredited users only. To fully exploit the MIP and raise potential interest from users, a new MIP instance (the Public MIP), that allows any scientist with an EBRAINS account to explore the system and the available features of the MIP, in a very simple way, was implemented.

2.3.1 The PUBLIC MIP

The Public-MIP[12] is a Virtual Machine installed on the EBRAINS RI and publicly available via EBRAINS accreditation on the EBRAINS website. With this new offer, EBRAINS users can explore the service, directly accessible via the EBRAINS MIP website. No installation, no download is required. Pending accreditation (personal, institutional), any person can analyse online all data placed in the public MIP and test the available MIP analytical tools. The public MIP currently only contains synthetic data, intended for exploring the system and testing its analytical capabilities.

What users can test and do on the Public MIP:

Experiment on synthetic datasets in Epilepsy, Dementia, F-Tract, Mental Health and TBI; select datasets, variables and co-variables to define the experimental parameters; choose between available algorithms (15 algorithms ranging from descriptive statistics, parametric statistics to supervised and unsupervised machine learning are built into the MIP); create, safe and share experiments to perform statistical analysis and last but not least access MIP Training Materials.

2.3.2 The MIP Hands-On Training Federation

A recent development was the creation of a Hands-on federation, which is a fully EBRAINS-hosted federation, implemented for the purpose of providing a suitable environment to run MIP training sessions.

The federation contains synthetic and public datasets in Epilepsy, Dementia, TBI and Mental Health. The test-run for this federation was in Dec 2021, during the EAN-EBRAINS Joint Workshop on the Future of Medical Data Sharing in Clinical Neurosciences (9-11 December 2021). With this training environment, the MIP teams will be able to respond in a very proactive way to user demands and they plan to provide user trainings on a regular basis.

2.4 Ongoing developments - current work on the MIP

2.4.1 New MIP Analysis Engine

To support the reinforcement of data privacy and safety guarantees of the MIP, ongoing work includes adding a Secure Multi Party Computation (SMPC) engine (see Figure 8). The SMPC subsystem engine serves as an extra layer of security and privacy. Computations performed locally on the hospital nodes are aggregated using SMPC into insights over the entire federated dataset. Data do not leave the hospitals unencrypted. Instead, data are secretly shared between the SMPC nodes using a secure importation protocol. A major challenge to overcome for meeting practical application requirements is scalability. We are experimenting with offering variable security levels, enabling a trade-off between security and speed. Further, we are working on offering a Differential Privacy (DP) guarantee through the injection of noise to computation results before publishing. Currently, extensive diagnostics and benchmarking of new implementations is performed on the SMPC engine, with the aim of finding the optimal balance between security and speed of execution.

[12] Public MIP https://hbpmip.link/

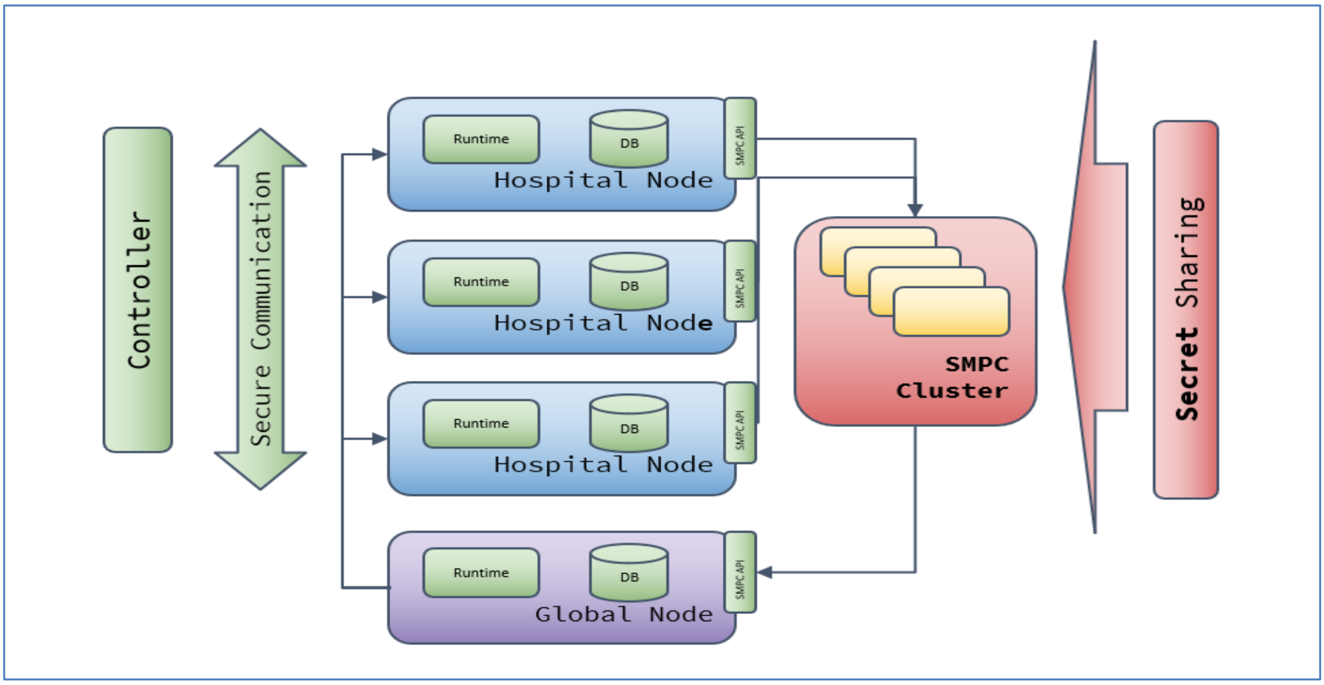

Figure 8: MIP high-level architecture with addition of the SMPC subsystem.

This work will result in the release of the MIP7.0, which will include added privacy guarantees for the MIP.

2.4.2 MIP New Federations and Collaborations

During the first 21 months of SGA3, new collaborations have emerged with the aim to create new federations:

The MIP entered a partnership with the scientific committee of the European Academy of Neurology (EAN) with the view to be used to federate datasets from many more neurological conditions. This was thoroughly discussed in a recent EAN-EBRAINS 3-day workshop, gathering more than 150 participants, which covered all neurological conditions and data-federation issues. As a first step towards this goal, an EAN use case has been launched, also supported by the European Stroke Organisation (ESO), which consists in federating national stroke registries from seven EU countries for a total of half-a-million patient's records. This use case, **the Stroke (FERES) federation,** is of particular interest since it involves the majority of EU stroke leaders who agree that such federation will enable addressing important research questions, which could not be tackled otherwise. Indeed, national stroke registries, which are used to certify all stroke centres from a given country and monitor their quality of care, are typically overseen by governmental bodies with no possibility to centralise the data with those from other registries outside their country of origin (see Table 1), which are not usually shared between countries.

Table 1: European registries interested in participating in the Stroke federation

Centre	City	Country	Registry
CHU Lille	Lille	France	French Stroke Registry
University of Basel and Basel Hospital	Basel	Switzerland	Swiss Stroke Registry
University of Basel and Basel Hospital	Basel	Switzerland	TRISP (Thrombolysis in Ischemic Patients cohort)
VASCage GmbH - Austrian stroke registry	Vienna	Austria	Austrian Stroke Registry
National and Kapodistrian University of Athens (NKUA)	Athens	Greece	European Stroke Registry
Irish National Audit of Stroke (INAS)	Dublin	Ireland	Irish National Stroke Registry
University of Aalborg	Aarhus	Denmark	Danish Stroke Registry
Azienda Ospedaliera di Perugia	Perugia	Italy	Italian Stroke Registry (IRETAS)

Preparations of legal agreements are underway; the MIP data managers are developing the data model in close collaboration with the leading PIs of each registry.

The NeuroCohort federation: the EPAD [13] (European Prevention of Alzheimer's Dementia Consortium) is closely working with the IMI Neuronet (Efficiently Networking European Neurodegeneration Research) [14] programme to expand and boost collaborations in research worldwide. In this framework, seven members of the former EPAD (see Table 2) implemented a pilot use case, federating synthetic dementia data with the MIP. Based on a positive evaluation of the pilot study, current goal is to federate up to 40 centres that will work on clinical data in the field of dementia.

Table 2: Centres participating in the NeuroCohort pilot federation

Centre	City	Country
BarcelonaBeta Brain Research Center	Barcelona	Spain
Fundación Cita-Alzheimer Fundazioa	Donostia-San Sebastian	Spain
University of Perugia, Centre for Memory Disturbances	Perugia	Italy
Swansea University/DPUK	Oxford	UK
University of Edinburgh	Edinburgh	UK
University of Montpellier	Montpellier	France
Instituto Neurológico Andaluz	Seville	Spain

Institutes in Barcelona, Donostia and Perugia have already started the deployment procedure, while the other centres are still completing the administrative steps. Both, the NeuroCohort data model and the synthetic cohort data are in preparation.

ADDIA Cohort: Amoneta Diagnostics SAS implemented a large Proof-of-Performance (PoP) clinical study in Alzheimer Diseases patients (the ADDIA study) to validate new blood biomarkers. Part of these data will be shared within the current MIP dementia federation. The legal agreements are signed and deployment is underway. The new data model is under discussion and will be prepared in collaboration with partners in the dementia federation.

The MIP joined two IMI (Innovative Medicines Initiative) projects, IMI SOPHIA[15] (Sep 2021) and IMI EPND[16] (Nov 2021) and prepared applications for a Horizon Europe call.

Epilepsy Federation: Deployment and update of the MIP in hospitals from European Reference Networks (ERNs), specifically Epicare, will start as soon as the Epilepsy Datamodel is finalised. An EBRAINS-embedded federation with 10 nodes holding synthetic data is already prepared for this purpose. Federated remote nodes will join this federation, as a result, the Epilepsy Federation will be a hybrid federation.

2.4.3 New MIP release - MIP6.5

Expanding the features and enhancements made available through the stable MIP6.4 release, the MIP6.5 release candidate brings additional capabilities to its end-users. It is currently tested and evaluated. The main feature of this version, among other enhancements and stability improvements, is the possibility to include additional variables in experiments, calculated in an ad-hoc fashion through user-defined formula expressions, while visualisation enhancements in core descriptive statistics calculations bring better insight to the user.

[13] EPAD https://ep-ad.org
[14] IMI Neuronet https://www.imi-neuronet.org/
[15] IMI Sophia https://imisophia.eu/
[16] IMI EPND https://epnd.org/

2.4.3.1 MIP User Interface

The front-end design is undergoing continuous improvements to enhance the user experience and satisfaction. We aim to maximise responsiveness, efficiency, and accessibility of the UI and further optimise the visual design of the solution.

Formula

We are implementing a new component for entering formulas on selected variables in the front end. Based on this formula, we enable end-users to include in experiments additional variables calculated in an ad-hoc fashion through user-defined expressions following well-defined and widely accepted and commonly integrated semantics, based on R-platform-available transformations. This functionality is used for Descriptive Statistics and Logistic Regression, and provides

- The ability to add transformations on continuous variables, with log, exp, centre and standardisation functions.
- A feature to add interactions between pairs of continuous variables.

This is a first iteration of the formula, which was designed with simplicity in mind, as parsing a handwritten formula is error prone for users. The forthcoming releases will see the integration of more functionalities (nominal variables transformation and interactions).

Statistics Insights

Further visualisation enhancements are made available in core descriptive statistics calculations, to bring better insight to the user on the available datasets.

Storybook

In the development environment, the front end now integrates "StoryBook[17]", a library which lists components in isolation. It will allow developers and data scientists to have a better overview of the visualisation library the MIP is offering, and to better shape the associated data types.

Upcoming version, MIP6.5 will contain:

- GraphQL implementation for Metadata, Descriptive Statistics
- GraphQL type definitions available for front end

2.4.3.2 The MIP Gateway

Starting with the MIP6.5 release, a new component – the MIP Gateway, situated between the front end and the federated analytics engine, is introduced in the system architecture. This new architectural element shifts the design and further technical development to a modular structure, largely improving the MIPs interoperability (see Figure 9) and flexibility to reply to user demands.

The MIP Gateway will largely enhance the MIP's interoperability aspects. Major improvements as compared to the current system, will be the ability to support multiple analytics engines, an agnostic front end for easier feature addition and the provision of a new stack of data/analysis visualisation outputs independent of the analysis engine.

Due to the gateway, the MIP will offer improved state management (i.e. management of the state of one or more user interface controls such as text fields, OK buttons, radio buttons, etc. in the graphical user interface) and separation of concerns (i.e. the software architecture design principle for separating an application into distinct sections).

More aspects of the MIP Gateway are:

- Generalised features and better testability
- GraphQL implementation with a comprehensive API Documentation and shared data structure

[17] StoryBook https://storybook.js.org/

These features will be introduced sequentially; the core Gateway mechanism, switching API calls from the front end to the Gateway, has already been implemented.

Technical documentation [MIP gateway](https://mip-front.gitbook.io/mip-gateway-doc/)[18]

In SGA3, major improvements have been made to the Medical Informatics Platform. Because of this, combined with major efforts in collaborations, broadening and enlarging the MIP user base, we expect to see the MIP community thrive and the take up of the MIP substantially increase. Overall, the MIP has turned into a more secure, robust, reliable and usable tool for health data federation, as illustrated by its recent attraction to a number of European consortia (see Section 2.4.2) and major academic organisations (European Academy of Neurology, European Stroke Organisation, European Reference Networks).

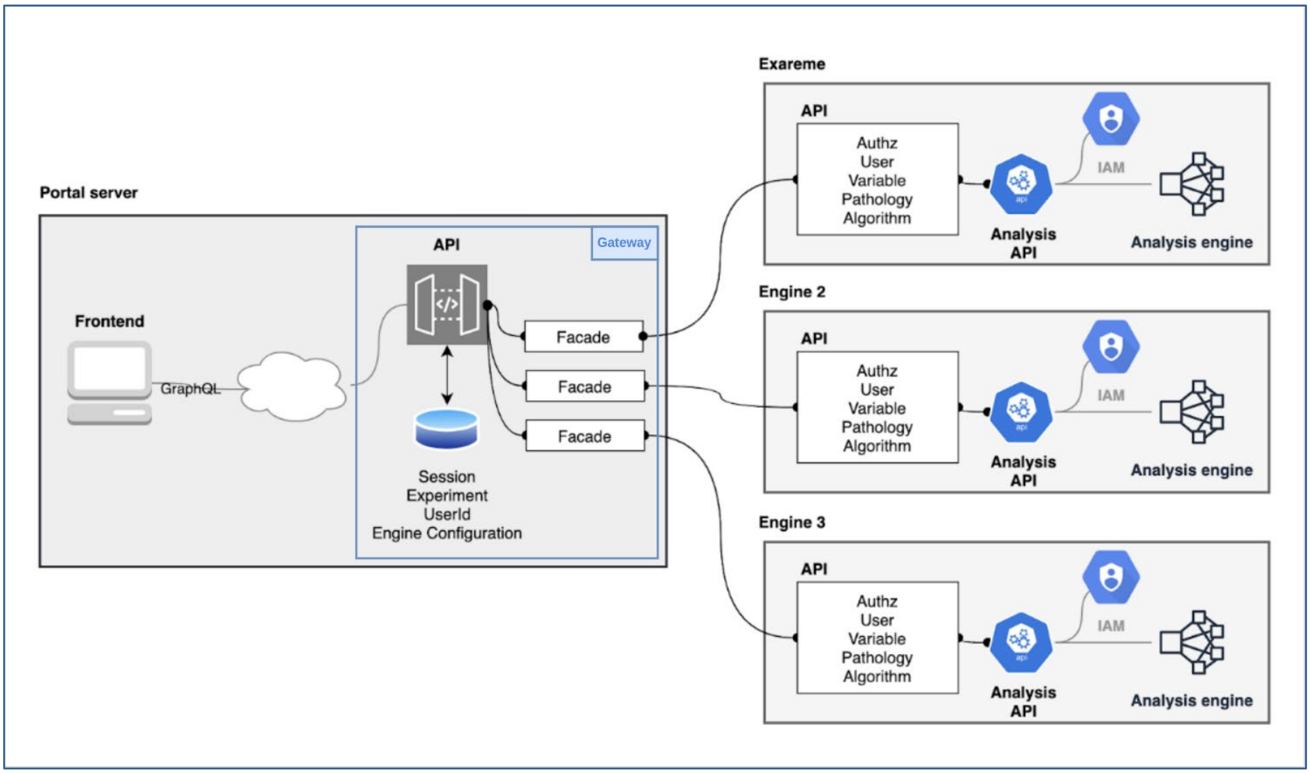

Figure 9: Modular MIP architecture with Gateway approach

The medical community, facing increasingly stringent data privacy regulations, starts showing interest to the unique value of data federation as an alternative to data centralisation. This is illustrated by an exponential production of scientific papers around the topic of “federated learning” or “federated analytics”, during the last two years. The previously cited paper on COVID-19 (Dayan et al. Nat Med 2021) truly represents a landmark in the field. Likewise, the recently launched EAN-ESO stroke use case is likely to become a game-changer for the MIP and its visibility. We are finalising the main publications of the dementia and TBI use-cases, while further expanding the ERN EpiCare federation on EBRAINS. We are very positive that an EAN-EBRAINS partnership will materialise by the end of SGA3, with the MIP becoming the central platform to federate clinical neurosciences data across neurological conditions.

Along these developments, the MIP will benefit from two major upgrades by the end of SGA3:

1) A reinforcement of its data privacy by integrating SMPC (secure multi-party computation) and differential privacy in its analytical engine (see Section 2.4.1)
2) Interoperability of its front end to connect to other analytical platforms

Through the above, the MIP will continue to stand as the most advanced and deployed free-of-use and open-source platform for clinical researchers to federate their health data. It has key advantages

[18] MIP Gateway https://mip-front.gitbook.io/mip-gateway-doc/

to be integrated in the future European Health Data Space (EHDS[19]), which is one of the priorities of the Commission in 2019-2025 and planned to promote better exchange and access to different types of health data to support primary and secondary use of data.

3. The Human intracerebral EEG platform - HIP

State of the art: Intracerebral EEG data is sparse, difficult to collect, and their processing or analysis requires domain and IT-specific knowledge and resources. Consequently, few standards, methodologies and platforms have emerged to permit the sharing or collaborative use of such data.

3.1 The Human intracerebral EEG data curation and management

3.1.1 The HIP Network

A large user survey sent to all 78 SEEG (Stereoelectroencephalography) centres known to operate in Europe, Asia and Oceania, established the interest of 95% of them in participating in a consolidated network of iEEG data providers in the development and use of the HIP (see Figure 10). Many of these centres are already involved in the HBP through the partnership established with the ERN EpiCARE network; several among the additional centres are part of the active SEEGMIP, EPINOV and F-TRACT consortia.

The management of the HIP data provider and user consortium will follow the HIP Consortium Charter, a document compiled to outline best practices and provide guidance for the collaboration, governance, rules for citation and publication, amongst other aspects. The charter is not considered a legal document, despite the recommendation to closely adhere to it.

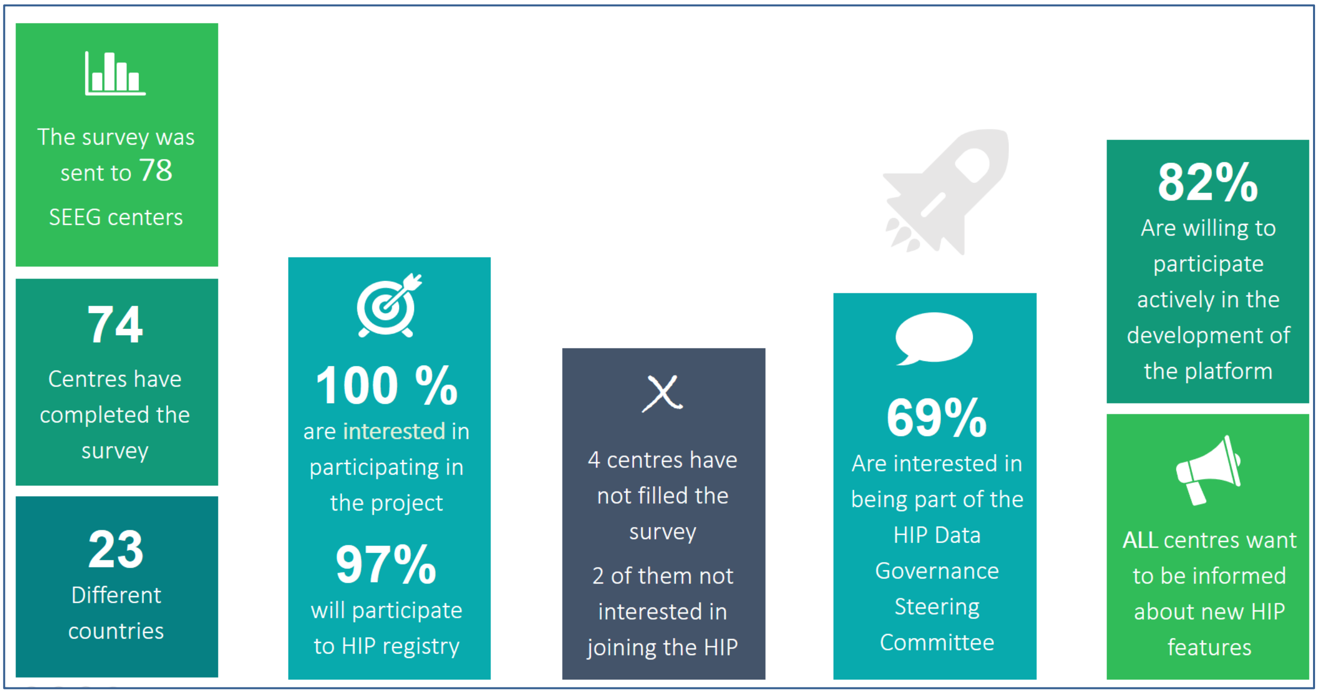

Figure 10: Results of the survey conducted with leading SEEG centres worldwide.

[19] EHDS https://ec.europa.eu/health/ehealth/dataspace_en

3.1.2 The HIP Legal and Ethical Framework

The HIP is an EBRAINS Service. Thus, several EBRAINS-related agreements need to be signed and accepted by users. Those are provided on the EBRAINS website, delineating the EBRAINS Terms of Use[20]. Various specific documents guarding the work on the HIP are under preparation, close to finalisation. The defined framework will enable a broad dissemination of the HIP to the scientific community and secure access and use of iEEG data.

The HIP registry: All HIP data will be included in the international HIP Registry coordinated by the CHUV Project Leader and Management Team, providing a sustained regulatory framework for collecting and storing these data. The purpose of this registry is to help delineate a multiscale model of the Human Brain, which is based on the large collection of available knowledge and data, fragmented across institutions and countries. The registry was produced in alignment with the Centre Opérationnel des Biobanques et Registres (COB) of CHUV, to comply with the CHUV legal and ethics rules. The registry is set-up for systematic data collection for research purposes and provides a first "layer" of compliance. It will include data collected during clinical practice and research data (typically from cognitive iEEG protocols) alike. While the registry does not need to be specifically ethics validated or approved by any other centre or country, the HIP will require all centres to get ethics approval from their local or national ethical bodies, to be allowed to collect and upload data to the HIP.

The registry comprises different types of data as described below:

- Retrospective AND prospective data
- Data acquired during clinical practice AND data acquired for research purposes

All the above include iEEG data, neuroimaging data used for the anatomical localisation of the recording iEEG leads, demographics and clinical data, as well as behavioural and cognitive data collected during cognitive tests.

All data collected in the HIP registry will be pseudonymised prior to any upload and covered by appropriate ethics and patients' consent. Pseudonymisation will include processing Brain MRI in order to remove all extra-cerebral tissue to avoid potential reconstruction of the patient's face. Prospective research data will be largely covered by the CogniEEG protocol, which has been specifically designed for the HIP. However, other iEEG research datasets, either prospective or retrospective, will be integrated in the HIP registry, provided that all legal and ethics requirements for data re-use and transfer on the HIP have been met (these might depend on the country of origin of the data). Likewise, iEEG data acquired during clinical practice will be available from institutions and patients with appropriate ethics clearance and legal authorisation for data re-use for research and transfer to another institution. In an increasing number of hospitals, this is covered through the so-called "general consent" forms signed by patients, authorising their clinical data to be re-used for research.

While the HIP registry is unique, it includes different media to store different data types, with the view to strengthen data privacy. iEEG and neuroimaging raw data will be stored in the private space of the HIP platform, where each participating centre has its own access-protected repository, where it uploads, curates and stores their data, and to which no other centre can get access. Behavioural and cognitive research data collected during iEEG cognitive tests will also be stored in the HIP private space.

In contrast, clinical data, providing pseudonymised information on patients' epilepsy and comorbidities, will be stored in a dedicated RedCap eCRF outside the HIP platform. In a majority of iEEG centres, this eCRF will be operated on the local RedCap of that centre. For the minority of iEEG centres, which do not have access to a RedCap server in their institution, the CHUV RedCap server will host their data. Yet, as for the HIP private space, each iEEG centre using the CHUV RedCap server, will collect and store its data in a password-protected centre-specific RedCap file, not available to other HIP centres.

[20] EBRAINS Terms of USE: https://ebrains.eu/terms/

This framework, physically separating the storage of raw data (iEEG, neuroimaging) from clinical data, aims at minimising the risks of data privacy breach.

Most importantly, the overall organisation of the HIP registry described above, entails the fact that all collected data in the HIP registry remain solely accessible for the data controller of the iEEG centre of origin. In other words, data stored in the HIP registry are harmonised across iEEG centres, but not shared between centres, as long as data are only collected and stored. Data sharing occurs at a later stage, when the HIP consortium agrees on performing multicentre analyses of a given dataset, available in the registry. A dedicated protocol is delineated and validated by the HIP consortium and submitted to the relevant ethics board(s). Once approved, selected iEEG data are pushed to the collaborative space of the HIP platform. Likewise, clinical data from the selected patients, which need to be used as covariables in the iEEG data analyses, will be extracted from their respective RedCap datasets, and sent to the HIP collaborative space. Once data analysis of a given iEEG protocol is finalised, all corresponding data can be removed from the collaborative space.

HIP CogniEEG Protocol: This protocol offers a generic framework to perform cognitive testing during iEEG, in order to generate new iEEG-based scientific knowledge on Human brain function and cognition. It includes a number of specific tests probing various cognitive domains, including working and episodic memory, executive functions, attention and various aspects of sensory integration. It is framed to seamlessly integrate new paradigms through amendments to the protocol. It is currently being finalised and will be submitted to the Ethical Committee in Vaud, Switzerland, before being submitted to all other participating iEEG centres and countries. The protocol includes a HIP-CogniEEG-specific informed consent.

CogEpiStim Protocol: In parallel to the CogniEEG protocol prepared by CHUV, in November 2021 CHUGA has validated a protocol of SEEG data re-use, allowing UGA, AMU, UCBL and CHUV to upload SEEG data on the HIP and to develop the HIP functionalities with such data, from the private up to the public space.

A **Data Transfer Agreement** (DTA) is currently being developed, covering the transfer of pseudonymised data to the HIP. It will be signed between a participating Centre and the CHUV prior to any data transfer. Data will be pseudonymised and/or anonymised prior to storage on the HIP platform. Once the draft DTA is finalised, it will first be approved by the CHUV Legal Department, then iterated/approved and signed by the institutions co-developing the HIP.

3.1.3 The HIP 3-tier architecture

The HIP Solution: The HIP offers a turnkey solution to scientists, physicians and researchers to securely curate, share and operate on intracerebral EEG data online. For its design, a strong focus was put on the use and integration of leading open-source software in the domain, to recombine those in a unique way, delivering the requested functionality within the given regulatory and legal framework. All applications, including desktop applications, are made accessible directly in the browser, anywhere, with no special IT requirements. The HIP allows programmatic, unattended use of applications that permit scripting or automation.

The benefits to its users and constituents are:

- First worldwide repository of its kind that supports visualisation, processing and analysis
- Delegated compute processing with a gain on performance and efficiency
- Increased exposure to application developers - no installation needed, apps can be tested or used without the additional burden of special IT resources
- Data curation and use of standards to foster interoperability and augment future data analytics potential
- Preloaded pipelines and workflows of the most adopted methodologies, helping users as scientific guides and enabling continuous automation support
- Increased international data interoperability, accessibility, availability and affordability of high-quality imagery, curation and analysis tools

- HIP primarily relies on BIDS as a common data format, but accommodates various other data formats

The platform high-level overview:

- Privacy-aware, distributed compute storage and remote application and display server environment to promote secure collaboration
- Compliant with ethics and international regulatory frameworks
- Cross-border sharing of pseudonymised visualised iEEG data, study and ethics protocols, analytical tools, workflows (all applications, including Desktop applications, run as in-browser-apps).
- Security and closed-circuit collaborative functions for data manipulation, analytics, visualisation and processing are preventing data exfiltration

The HIP encompasses all aspects of the three-tier architecture methodology in its solution as a platform as a service. The user-experience was a major driver for the development and integration strategy. Therefore, both the client and server-side applications, tools and services, were adapted to be executed and operated on remote server instances.

The choice has been to pick an innovative approach, and reconcile the storage, compute and desktop experience over browser-based Web applications, while retaining compatibility with standard data transfer protocols. Through its modular and interconnected architecture, resources can be hosted independently in different cloud and HPC environments, and be addressed programmatically via the HIP API functions. The design permits the delocalisation of traditional Desktop and UI applications to be executed and operated on remote server resources, leveraging both excellent CPU and GPU performance.

The HIP offers an innovative privacy-enabling data-sharing ecosystem with three defined spaces (see Figure 11):

Personal Space: a secured, access-controlled, private and dedicated space for data providers to upload, store and process their own pseudonymised or anonymised data. Data Green Room and entry-point to the HIP.

Collaborative ("collab") Space: an access-restricted space, where curated and pseudonymised data can explicitly be shared with other accredited HIP Users for operating a specific and previously approved data analysis. The decision to move data from the Private Space to the Collaborative Space, for collaborative use, can only be taken by the Data Controller of that Personal Space. The Data Controller also agrees to the research studies and publications to which their data can contribute in the Collaborative Space.

Public Space: a space for public, fully anonymised iEEG data made available to any scientist, findable in the EBRAINS Knowledge Graph. Pursuant to the EBRAINS Access Policy, access to such data shall not require an EBRAINS accreditation. Decision to move subsets of data from the Personal or Collaborative Space to the Public Space can only be taken by the Data Controller of the corresponding data, provided that the data has been anonymised and patient's informed consent to make their data public, was provided.

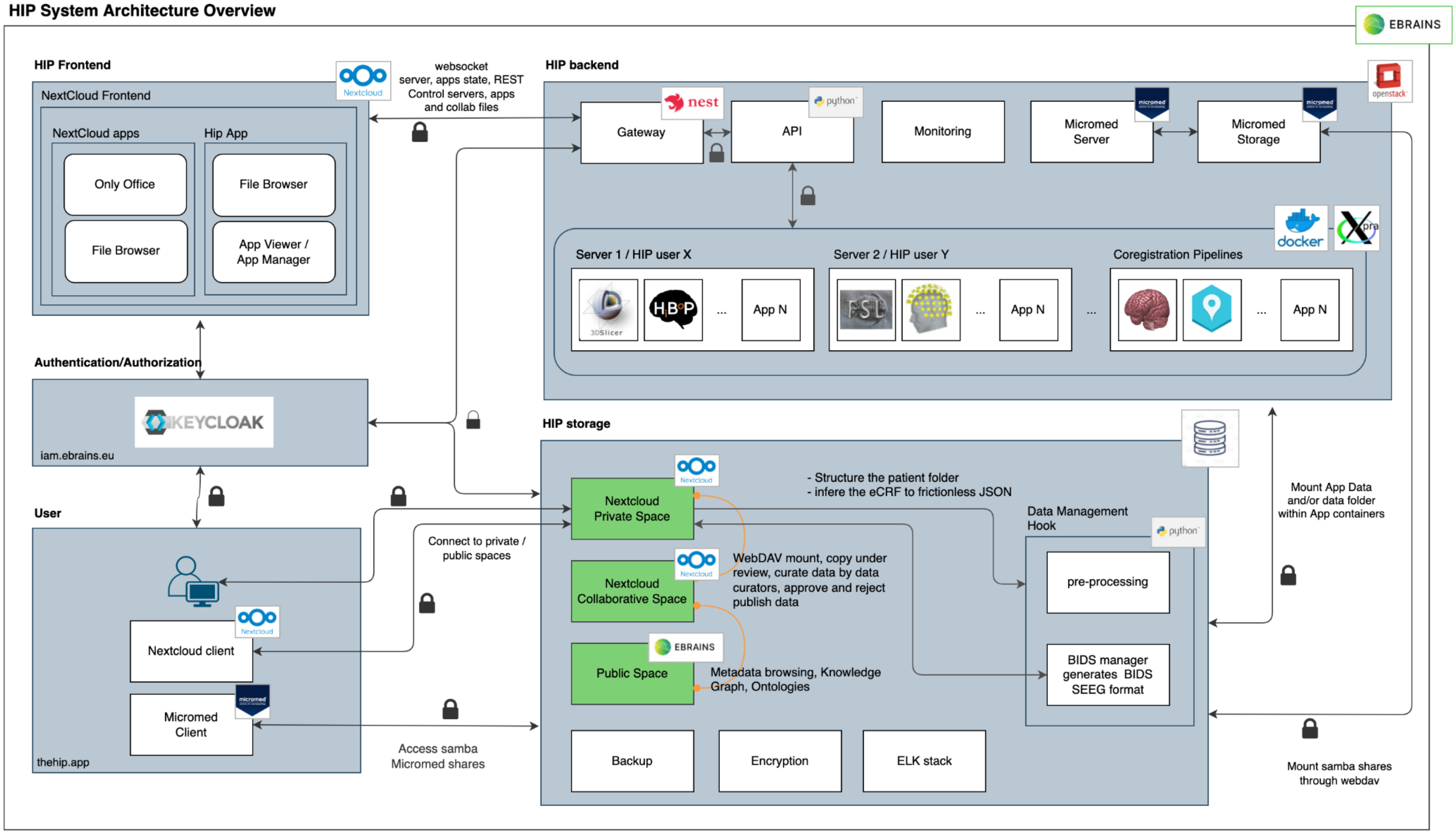

Figure 11: The HIP system architecture - a high-level overview

3.1.4 HIP Storage and Compute

The HIP is designed to provide a High Performance Compute and storage environment with horizontally distributed components; security and data safety are predominant in its conception. **Key decisions around the design of the storage and compute solutions are yet to be taken, best solutions are being investigated.** To address the risks of data breach or data loss, the HIP divides the different spaces into separate storage instances that will be federated. Such federation capabilities can permit the distribution of multiple instances for each space, and provide data backup and system redundancy. The HBP offers, as part of the Interactive Computing E-Infrastructure project, access to compute and storage resources of the Fenix infrastructure[21], a network of five European supercomputing centres. The HIP enables collaboration via standardised File Transfer Protocols and the use of Virtual Desktop Infrastructure environments. VPN and other networking technologies take care of the secure transport of data.

Currently, it is located on the Swiss Supercomputing Centre CSCS, and, for a pilot-project, on Exoscale[22], a European provider for secure cloud infrastructure services, situated in Switzerland. Thus, currently, the infrastructure hosting the HIP prototype is located in Switzerland.

Storage and Compute:

- The storage back ends will be a collection of federated nodes, allowing for easier growth / expansion, minimising data corruption risks
- Implementation of Kubernetes container-orchestration system
- Storage facilities can be in multiple countries and data centres to offer adaptive ethics and legal compliancy
- Compute operations can be localised to specific regions for ethics compliancy

3.1.5 The HIP standardised procedures and related tutorials

Major development efforts are made towards providing App Deployment Frameworks based on Docker and other containerisation technologies for the integration of applications and services, made available by the HIP development teams, and in the long-run also for future co-development of the platform with external users.

HIP user documentation is being designed to assist end users in accessing the platform and operating all services. With this perspective, it will provide comprehensive step-by-step guides and video tutorials, instruction manuals, frequently asked questions (FAQs) and troubleshooting sections.

The documentation will be web-based and available on an open source and long-term sustainable documentation hosting provider: readthedocs.org, redmine.org and ebrains.eu are currently being investigated in this respect. Work on the documentation is ongoing and will evolve with the maturity of the platform; the HIP user documentation[23] is drafted and hosted on a github repository.

The documentation will consist of:

- Starter guides to help the user getting access to the platform. This includes tutorials for the creation of a HIP account with relevant accreditations and an emphasis on the data controller responsibilities and FAIR data principles. How to connect to the HIP portal and how to access its various services (data uploading and storage, BIDS conversion tool, private and collaborative spaces, app library and processing pipelines) and internal tools (wiki, support/assistance)
- Data storage and sharing guides to help the user upload and share data on the platform. How to prepare the data (eligibility, pseudonymisation, curation and uploading conditions). How to

[21] Fenix RI: https://fenix-ri.eu/
[22] Exoscale: https://www.exoscale.com/
[23] HIP User Documentation: https://github.com/manikbh/HIP-doc

upload the data using dedicated tools (web-based uploader, file transfer protocol such as WebDav or dedicated client software for Micromed data). How to harmonise data according to BIDS standard using the integrated BIDS conversion tool. How to transfer a validated BIDS database into the collaborative space and share it with other HIP users inside a working group associated with a research protocol.

- Data analysis guides and use-case tutorials to help the user explore and process data and share results. This includes guides to initiate private or collaborative working sessions and start apps / pipelines available in the HIP software library. How to share analytical tools and results. The HIP is built on top of pre-existing iEEG tools coming with their own documentation that will be made available via integrated links.

3.2 The Human intracerebral EEG data and analytical tools integration

The HIP offers a dedicated software library (Section 3.2.1), ready to use and to be instantiated on remote cloud infrastructures. Each app lives in its own docker container and has direct access to the GPU of the host machine. Collaborators can access the integrated apps through sessions on the HIP via their web browser.

The HIP supports the entire field of iEEG-based research; i.e. multiscale investigations of cognition, consciousness, connectomics and related disorders. It provides the software infrastructure to prepare and process iEEG data to achieve the scientific objectives requiring the use of human intracranial EEG (iEEG) data, in the Human Brain Project mainly WP1 (The human multiscale brain connectome and its variability – from synapses to large-scale networks and function) and WP2 (Networks underlying brain cognition and consciousness). During SGA2, work to implement an infrastructure to manage and process iEEG data has been started, but was not fully integrated in the historical MIP. It required the specification of the BIDS (Brain Imaging Data Structure) for iEEG (BIDS-iEEG) format (international consortium involving UGA) (Holdgraf et al., Sci Data, 2019)[24], the development of a BIDS-iEEG data manager led by AMU (Roehri et al., Neuroinformatics, 2021)[25] and the generation of a BIDS-iEEG database (CCEP database accessible through ebains.eu, led by UGA). Building on the preliminary work performed in SGA2, we integrated a suite of analytical tools for iEEG data analysis in the HIP platform.

The current prototype platform is regularly tested in the end-user mode and the development mode. Testers can send feedback straight from their beta version, simply by taking a screenshot and sharing detailed feedback with a crash report immediately after a crash occurs. Testers and developers convene in daily conference sessions to assess the development and integration progress.

3.2.1 *The HIP Library - In-Browser Apps*

To date, a comprehensive, but still growing, set of tools was packaged as In-Browser-Apps and included in the HIP.

Table 3: The HIP Library - In-Browser Apps

Software/Tool integrated	Short Description/Link
AnyWave	Matlab-based software made for processing EEG, MEG, ECoG and SEEG. Developed by AMU (HBP, France) https://meg.univ-amu.fr/wiki/AnyWave
Brainstorm	Matlab-based software made for processing EEG, MEG, ECoG and SEEG. Developed by University of South Carolina (USA) and Montreal Neurological Institute (Canada). https://neuroimage.usc.edu/brainstorm/Introduction

[24] https://doi.org/10.1038/s41597-019-0105-7
[25] https://doi.org/10.1007/S12021-020-09503-6

Localizer	Application to process raw brain recordings to obtain the frequency responses using a combination of band pass filter and Hilbert Envelope. https://github.com/CRNL-Eduwell/Localizer
HiBoP	Unity-based software made for visualisation of SEEG data. Developed by UCBL (HBP, France). https://github.com/hbp-HiBoP/HiBoP
Slicer	Community-based software for imaging visualisation. https://www.slicer.org/
FSL	Software made for processing of MRI, FMRI and DTI data. Developed by Oxford University (UK). https://fsl.fmrib.ox.ac.uk/fsl/fslwiki
MRIcroGL	Compiled software made for the visualisation of MRI. Developed by University of South Carolina (USA). https://www.nitrc.org/projects/mricrogl/
BIDS_Manager	Python-based software made for creating and managing BIDS-iEEG databases: Developed by AMU (HBP, France). https://github.com/Dynamap/BIDS_Manager
Micromed	Software to review SEEG recordings. https://micromedgroup.com/download/systemplus-evolution-user-manual/ SystemPlus EVOLUTION™ User Manual - Micromed Group
dcm2niix	A script to convert MRI data. GitHub - rordenlab/dcm2niix: dcm2nii DICOM to NIfTI converter: https://github.com/rordenlab/dcm2niix
Freesurfer	Software made for processing of MRI data. Developed by Harvard University (USA). https://surfer.nmr.mgh.harvard.edu
Under Preparation for integration	
CCEP Manager	https://ccepmanager.readthedocs.io
IntrAnat	https://f-tract.eu/software/intranat
TVB-HIP-APP	https://github.com/ins-amu/hip-tvb-app

Each piece of software listed in Table 3, was installed on the HIP and meticulously tested. The performance of each piece of software was evaluated and the feedback from the end-user and developers of the tools allowed adjustments of the additional plugins and adapt the functionality according to users' requirements. Each new compilation required de-novo testing, which allowed detection and fixes of eventual glitches.

A demo session to present the current state of the HIP was organised for the HBP Summit Science Market in October 2021.

3.2.2 The integration of BIDS-iEEG manager software tools

The HIP integrates in its library (see Section 3.2.1) several software tools (as example see Figure 13), allowing to organise the data (SEEG or anatomical imaging) into the BIDS-iEEG folder architecture, to process them (SEEG or anatomical imaging), and to visualise the results.

As implemented on the HIP prototype, the HIP user can see the available apps in a tabular form (Figure 12). Each app can be opened in a single or in multiple sessions. In addition, several apps can be opened in a single session.

The available apps are mainly open-source software that can read data stored in the BIDS format. Equally, the HIP manages upload and data import of other data formats, as defined by the apps. The different apps allow to design multiple scenarios for the processing of iEEG data. They are made interoperable through the use of the BIDS formalism, which allows to store raw data in a standardised way, and to centralise the main outputs of the different apps in a dedicated place (derivatives folder of BIDS). However, most apps have their own database on the HIP, to avoid losing relevant information.

The BIDS data format is thus critical to facilitate data storage, data re-use and data sharing. To manage this important aspect, AMU has developed a dedicated tool called the BIDS Manager that has been installed on the HIP. The BIDS Manager has been used to generate BIDS-iEEG datasets provided to the EBRAINS Knowledge Graph (CCEP database developed by UGA[26]), which is being prepared for testing the HIP (COGEPISTIM trial by CHUGA, initiated in November 2021).

3.2.3 *Integration of visualisation tools*

An innovative solution, i.e. 3D rendering in the HIP virtual environments, was developed by the CHUV team. Now, visualisation of data for data quality control and data preparation, or for exploring the results of the processing, is performed using the HIP apps as if they were installed in a local environment.

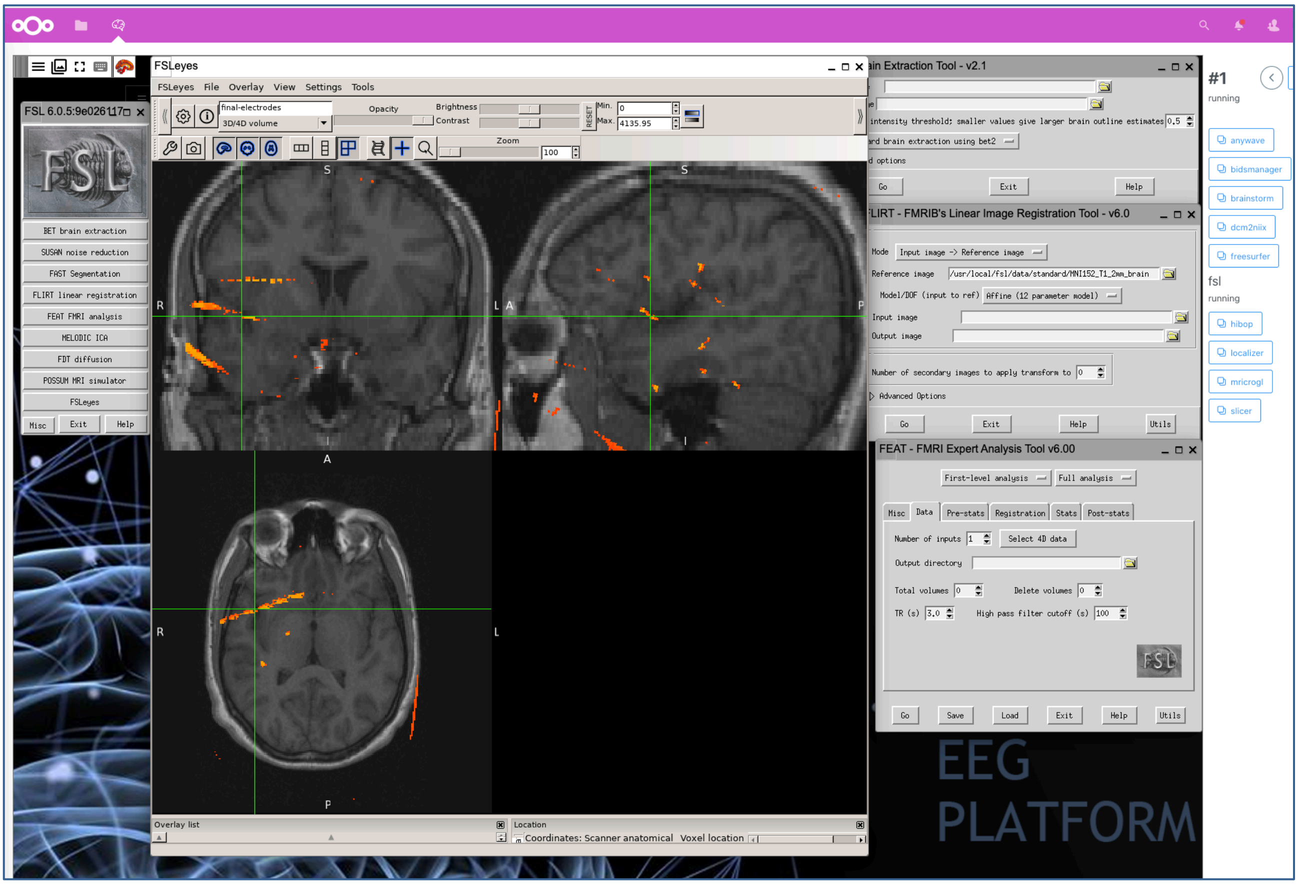

Figure 12: A HIP session, running FSL, displaying other available apps in the sidebar

[26] CCEP Data Base https://search.kg.ebrains.eu/instances/Dataset/ebe50517-41d5-4029-9355-04f1e49e23c8

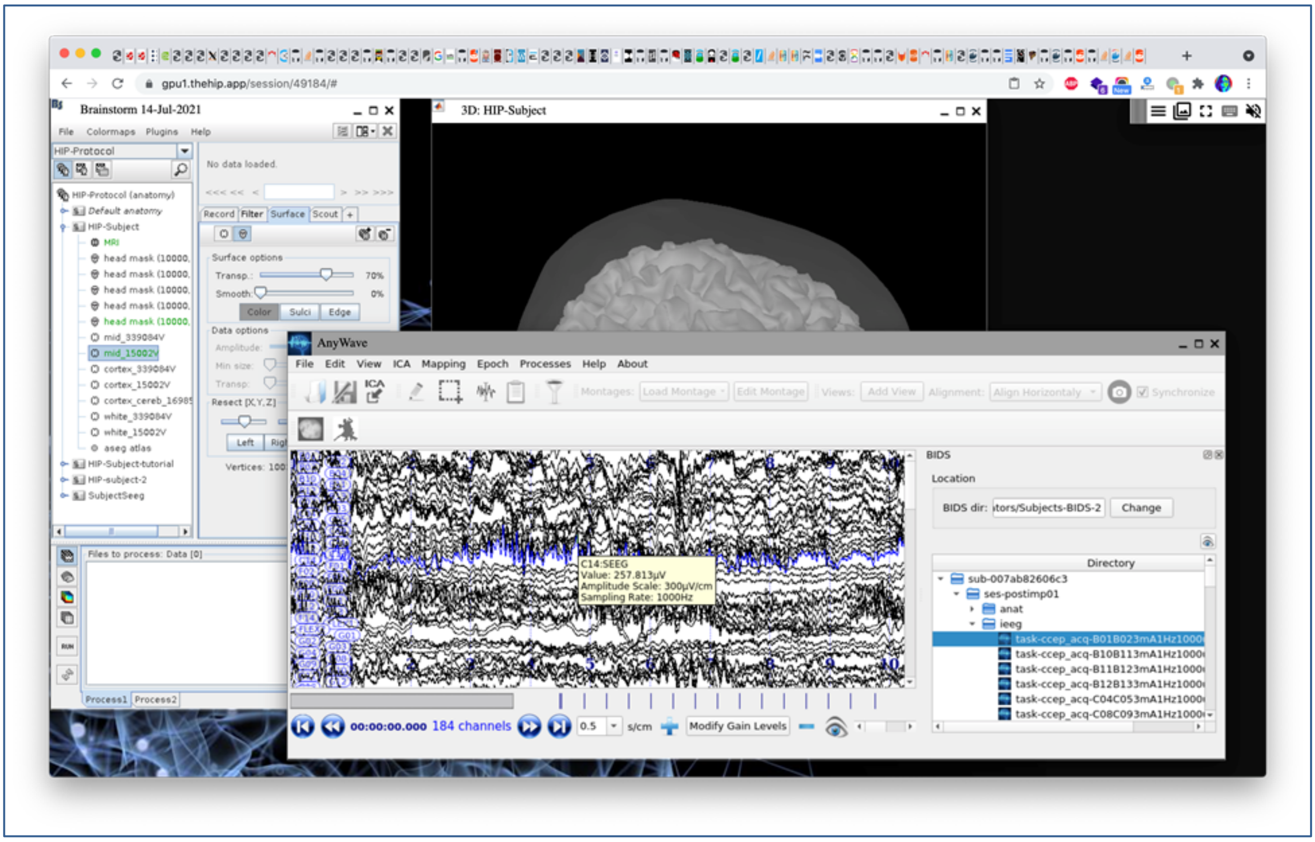

Figure 13: Anywave and Brainstorm in-browser apps opened in a single session on the HIP

3D rendering is required for positioning SEEG electrodes in the anatomy of subjects (e.g. using 3D sclicer), for reviewing SEEG traces (e.g. using Anywave), for assessing seizure networks (e.g. using Brainstorm) or brain connectivity (e.g. using HiBoP, Figure 14), just to name a few relevant use cases.

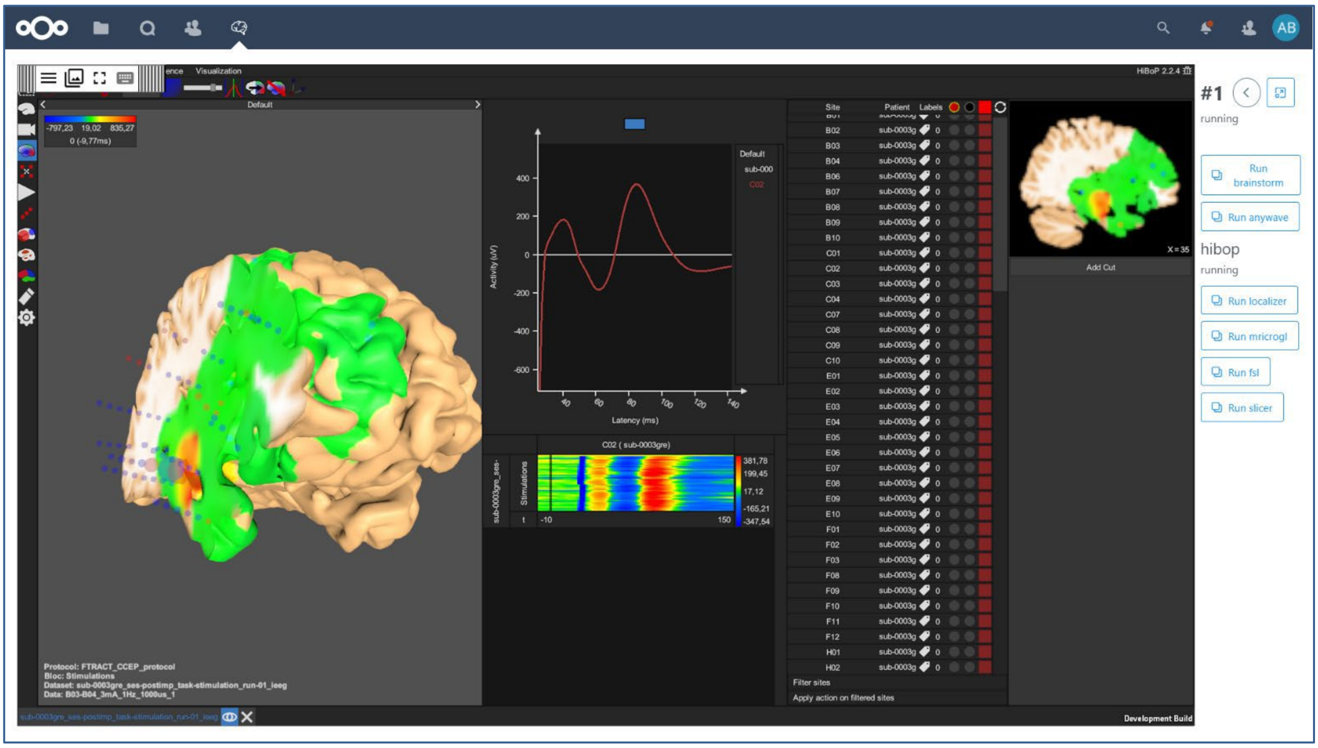

Figure 14: Data visualisation on the HIP: Use of the HiBoP app for 3D visualisation of evoked potentials

3.3 The Human intracerebral EEG data services: integration with other EBRAINS Services

A versatile in-browser app is being developed to support TVB structural and functional workflows on the HIP platform, schematised in Figure 15. This app incorporates several third-party libraries such as Siibra for Knowledge Graph and atlas access, PyUnicore for access to HPC resources and BIDS compatibility with TVB to process HIP data with the TVB. It presents a set of user-friendly Jupyter notebooks, which can be used with little to no need for programming, yet exposing many of the clinically relevant use cases available with TVB on the HIP. This app is currently under development for deployment on the HIP platform, it also enables interoperability with the wider EBRAINS platform.

The TVB-HIP app was designed to support workflows required by WP1 Showcase 2 (Improving epilepsy surgery with the Virtual BigBrain), which will make use of the HIP-TVB app to provide a complete workflow from raw clinical SEEG & MRI data to clinical report of seizure source estimates, integrating individual and high-resolution data. Key steps in this workflow include the ability to adapt high-resolution data from the brain atlas services to improve the modelling of an individual patient, and use of high-performance computing infrastructure.

3.3.1 The HIP structural and functional processing workflows

The integration of HIP workflows with TVB, Big Brain atlases, EBRAINS Knowledge Graph, HPC services and related software in HIP will enable a seamless translation of patient-specific information (data) into patient-specific models (parameters), allowing the creation of individual virtual brains, capable to serve as template for target fitting (for instance, epileptogenic zone) or in-silico clinical hypothesis testing. The in-browser TVB-HIP app enables these steps by providing interop workflows with these EBRAINS services.

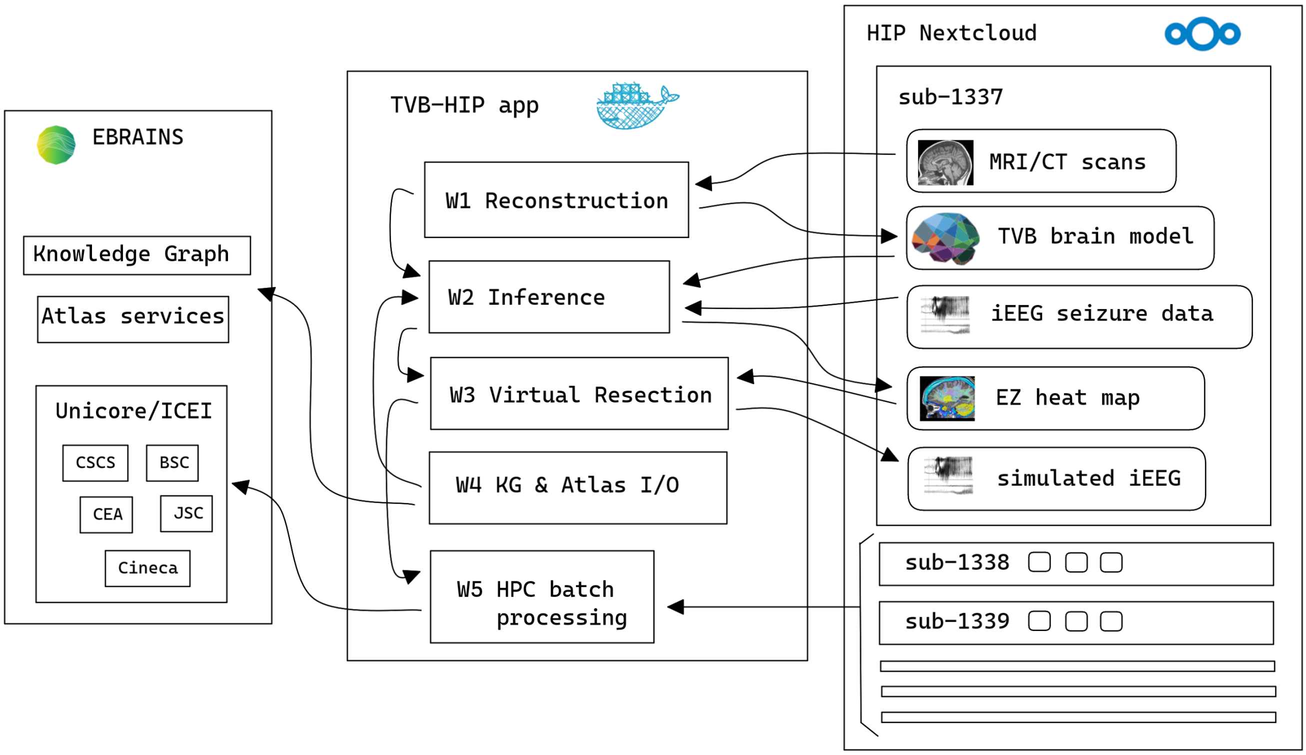

Figure 15: TVB-HIP in browser app and its relation to other components and data flows.

In addition to TVB-related activities, UGA (now AMU) has initiated the process of feeding the Atlas services (SC2) with SEEG data as if processed on the HIP using the CCEP manager. These data represent the F-TRACT atlas (f-tract.eu), which provides from CCEP unique information on longrange fibres' directionality and dynamics. These data are integrated using the OpenMinds framework of the EBRAINS Knowledge Graph. More details in Section 3.3.5.

3.3.2 TVB workflows and pipelines

The HIP-TVB integration will be supported by facilitating and optimising access to HIP data, as well as by enabling its pre-processing and rendering in the same, shared reference space. Co-registration of structural data enable applying TVB workflows from seizure simulation for virtual resection studies to inference of model parameters representing epileptogenic zones. The in-browser TVB-HIP app enables these steps by providing all key software requirements in set turn-key workflows, which also interoperate with HPC services to offload workflow steps too computationally intensive for the HIP services.

3.3.3 HIP Adaptors

The TVB-HIP app should interoperate seamless with other software in the HIP ecosystem. A major role in this interoperation is the construction of adaptors which import or export data in compatible formats. Native TVB support for BIDS format is currently being implemented by the TVB team, with whom we are coordinating the HIP use case. We have import functionality for Gardel electrodes, partial support for Brainstorm files; the latter will be improved once we have concrete datasets produced from Brainstorm. The set of tools oriented around iEEG time-series data, because TVB-HIP makes use of MNE-Python, can already import data converted by AnyWave, and export BrainVision-format data, and ongoing use cases may result in implementation of other formats as well. Finally, for data such as parameter estimates that are more suited to a functional anatomical overlay, we have implemented export in Nifti formats, which can subsequently be converted to any required format. To the extent possible, we adhere to standard coordinate systems such as MNI, allowing interoperability with brain atlasing tools.

3.3.4 The HIP Access, Use and Accreditation Modalities

Through its implementation and use of Nextcloud[27] as data storage solution, the HIP offers industry-leading security capabilities (Health Insurance Portability and Accountability Act (HIPAA) and HITECH compliant) and is fully GDPR compliant. Nextcloud supports File Access Control, built-in monitoring and full auditing capabilities. It uses industry-standard SSL/TLS encryption for data in transfer. Data at rest in storage can be encrypted using a default military grade AES-256 encryption with server-based or custom key management. Optionally, and on a per-folder base, data can be end-to-end encrypted on the client, with the server assisting in sharing and key management using a Zero-Knowledge model.

The HIP needs to comply with national and international laws and regulations, comprising principles such as intellectual property rights and the protection of privacy, ethical considerations and security regulations (Section 3.1.2), when designing rules and conditions for Access and the use of the platform in any of the aforementioned spaces (Section 3.1.3). Only authorised users will have access to the HIP. In an initial phase, accreditations are managed by the HIP Leadership (CHUV and AMU). Once the HIP is in full use, oversight will be granted to the Data Governance Steering Committee, as outlined in the HIP Charter. Users must request an EBRAINS account by contacting support@ebrains.eu and specify if they are member of one or multiple working groups in the HIP. EBRAINS KeyCloak Authentication for single sign-on and the fine-grained accreditation process are being designed; several key aspects are already described in the HIP Charter; now the technical design and implementation as well as alignment with EBRAINS are being pursued. The HIP Prototype platform is not ready for onboarding of external users and their pseudonymised data yet.

[27] Nextcloud https://nextcloud.com/industries/healthcare/

3.3.5 Integration of F-Tract in the Multilevel Human Brain Atlas

The data collected in the F-Tract ERC project integrate SEEG recordings from around one thousand drug-resistant epileptic patients. A subset of these data is already integrated into the EBRAINS Knowledge Graph (EBRAINS - CCEP database of the Medical Informatics Platform[28]; DOI: 10.25493/SV5Z-FSB). The development of the HIP, as the dedicated platform, supporting collection, storage, analysis and sharing of such data is still ongoing and will in the future further leverage the SEEG dataset in the HBP context as described below.

First, the SEEG-derived human brain connectivity maps will be published in the EBRAINS Big Brain Atlas. These maps are prepared and are available for the most widely used brain parcellations, including those used by the Big Brain Atlas (Juelich atlas and Automated Anatomical Labeling (AAL) atlas); data upload to EBRAINS is in progress, subject to continuous iterations with the EBRAINS Data Curation team. Current work implies the use of the OpenMINDS metadata model, to include F-TRACT data and metadata in the EBRAINS Knowledge Graph, allowing users to find the data by e.g. keywords or other links defined by the graph.

Second, the above-mentioned connectivity maps will be integrated into the online Multilevel Human Atlas Viewer[29] available on EBRAINS. Not only connectivity patterns provided by the SEEG-derived probabilistic atlas will be available in the viewer, a user will also be able to browse properties of connections between given parcels. An example of such property is speed of the neuronal signal transmission. The F-TRACT data descriptor, which will be displayed in the Knowledge Graph, was submitted to the HBP Curation Team, the unit responsible for integration into the Atlas Viewer, for analysis.

The work described above is executed in collaboration with the EBRAINS Atlas Services and EBRAINS Data and Knowledge Services, covering the technical steps necessary to deliver the SEEG-derived maps on EBRAINS to end-users. Additional efforts of WP1 (The human multiscale brain connectome and its variability – from synapses to large-scale networks and function), are the integration of connectivity maps obtained with two different modalities (collaboration with CEA): SEEG and diffusion MRI. Consistency between the two modalities is studied and quantities inaccessible from each modality alone, such as spatially extended signal propagation speed, are computed. Current work focuses on removal of extreme values from the analysis and on the presentation of results.

3.3.6 Integration of new HIP apps and pipelines

Several apps are being prepared for installation on the HIP (see also Section 3.3.3), such as IntrAnat for electrode localization[30] and the CCEP manager[31], which is a standalone implementation of the F-Tract analysis pipeline allowing the processing of iEEG data coupled with electrical stimulation for the extraction of Cortico-Cortical Evoked Potentials (CCEP) features and the generation of connectivity atlases (developed by AMU).

Specific use cases (e.g. SEEG electrodes localisation, epileptic seizure analysis, connectivity analysis from direct electrical stimulation) are being designed and will be fully documented, when the HIP reaches its production phase. As an advanced use case, AMU-UGA have been working on re-implementing the complete pipeline for processing the F-TRACT[32] database of cortico-cortical evoked potentials with a dedicated BIDS-app, called the CCEP manager, that will be installed on the HIP.

With the integration of a large number of sophisticated tools and workflows that require an orchestration of apps, the demand on the user interface is growing. The current prototype UI will

---

[28] CCEO database https://search.kg.ebrains.eu/instances/Dataset/ebe50517-41d5-4029-9355-04f1e49e23c8
[29] Multilevel Human Brain Atlas https://ebrains.eu/service/human-brain-atlas/
[30] IntrAnat https://f-tract.eu/software/intranat
[31] CCEP Manager https://ccepmanager.readthedocs.io
[32] F-tract Functional Brain Tractography Project – https://f-tract.eu/

evolve to an advanced HIP User Interface, informed and guarded through co-development with end-users to meet their requirements. One imminent use case is an improved or tailored data upload facility for conversion of data to the BIDS format.

Ongoing work to facilitate planned activities:

- Framework development for supporting auto-deployment of apps-in-browser for development teams and, once in production, for onboarding of external tools, brought in by end-users
- Preparation of a base image for a JupyterLab desktop app, as required e.g. for TVB integrations
- Optimisation of file transfer throughput to load data in remote apps
- BIDS manager conversion (Common Workflow Language, CWL)

3.4 The HIP Prototype – platform status in December 2021

The HIP is a prototype platform in December 2021. Many aspects of the architecture and desired functionalities are defined, as outlined in the sections before, several aspects are under development, some components and connectors still need to be designed, some key design decisions are currently being made. To put the development progress of the HIP in context: at the beginning of SGA3 (April 2020), no other HBP service was handling or managing sensitive clinical data. Neither the infrastructure requirements or security provisions for this kind of data, nor the legal framework were in place. Now, 21 months later, we have realised an innovative prototype platform, mastered many challenges, and a variety of components and workflows are ready for integration. The coming six months will lead to a substantial leap in the HIP, lifting the prototype to a production platform, enabling first users to upload their pseudonymised data and test the platform functionalities. This next phase will be critical for testing and refining the implemented concepts with end-users and stakeholders involved and refine or further develop the HIP concept.

Below, a mere technical summary of the current status is shown – all other aspects were already depicted in detail in earlier sections:

3.4.1 The HIP repositories and Back end

Repositories:

- All git repositories on github with auto-sync on gitlab instance on EBRAINS CSCS, Docker repository on gitlab

Back end:

- Back end CI: Automatic build of docker images on push; App-in-browser
- GPU/CPU integration on EBRAINS CSCS
  - Framework based on docker; Server: xpra; Process manager: pm2; Webserver: caddy
  - Base images: vgl, matlab-runtime, webdav, terminal
  - Dockerfile template available for easy app integration by partners

3.4.2 The HIP Front end and Gateway

HIP Front end: Prototype Front end

- CI: Automatic build of Nextcloud app and API; CD: *to be implemented*
  - Nextcloud Private Space: HIP app Session and App control (Mono repo service under React; User tests)

- NextCloud Collaboration installation (*to be implemented*)
- File manager for browsing collab data without disclosing data (*to be implemented*)

HIP Gateway: allows programmatic construction of workflows, application execution and management of server resources.

- Hub for Services APIs communication; allows communication among various apps, resources, components; programmatically allows selection of datasets and makes them available at the appropriate workflow step (*to be designed*)
- State Machine to enforce correctness of sessions/apps sequences (start, stop, destroy)
- User state (sessions, apps, credentials) in Redis (messaging and caching service)

4. Looking Forward

4.1 MIP future developments

- Implementation of new features: Of immediate benefit for users will be the download option for aggregated results of an experiment (jpeg, png, pdf, svg vector image) and likewise the integration of the MIP Data Quality Tool as a web interface in order to harmonise local data.
- Enhancement and enforcement security over the whole cloud infrastructure with SSO (Single Sign-On) and deployment of the full security solution on all servers.
- Addition of privacy guarantees with the use of Secure Multi Party Computation (SMPC) and Differential Privacy (DP).
- Release of a MIP local version to be downloaded from EBRAINS and easily installed together with its dedicated virtual machine by end-users on their own professional laptop.
- Interoperability of data with other open standards through tight integration of MIP data management and workflows with the EBRAINS Knowledge Graph, which is promoting the open Metadata Initiative openMINDS. Development of ingestion pipelines, APIs and standardisation workflows between different data formats to ensure a seamless exchange between the two EBRAINS services and allow appropriate tagging, annotation of the sensitive clinical data in the MIP, protecting embargoed information from exposure.

4.2 HIP future developments

- Security: The prospective integration with the EBRAINS service for sensitive data, the Health Data Cloud (HDS)[33] will determine many of the security, data safety and high availability modalities, as required for an EBRAINS integrated service. This new EBRAINS service started its development in the beginning of the year and will lead to synergies, benefitting both the HDS and HIP in their maturation.

[33] Health Data Cloud (HDS) https://www.healthdatacloud.eu/

5. ANNEX

Table 4: Documentation and other references

Technical Documentation	
Main MIP Technical Documentation Repository	https://github.com/HBPMedical/mip-docs
MIP Deployment	https://github.com/HBPMedical/mip-deployment/blob/master/doc/Readme.md
MIP Federated Algorithms	https://github.com/madgik/exareme/tree/22.2.3/Exareme-Docker/src/mip-algorithms
MIP Gateway	https://mip-front.gitbook.io/mip-gateway-doc/
Other relevant links for the MIP	
The Medical Informatics Platform (MIP) - website	https://ebrains.eu/service/medical-informatics-platform/
The PUBLIC MIP	https://hbpmip.link/
Redolfi A, et al., 2020: Medical Informatics Platform (MIP): A Pilot Study Across Clinical Italian Cohorts. Front. Neurol. 11:1021. **P2409**	https://doi.org/10.3389/fneur.2020.01021
The Medical Informatics Platform (MIP): Analysis and options for its exploitation in the healthcare market	https://sos-ch-dk-2.exo.io/public-website-production/filer_public/82/4a/824ac0ed-d2f6-495f-b560-c5647838c46a/upm_hbp_mip_final_dt.pdf
IMI Neuronet (Efficiently Networking European Neurodegeneration Research)	https://www.imi-neuronet.org/
Vagrant	https://www.vagrantup.com/
DEBIAN GNU	https://www.debian.org/
EPAD (European Prevention of Alzheimer's Dementia consortium)	https://ep-ad.org
IMI Sophia	https://www.imi.europa.eu/projects-results/project-factsheets/sophia
IMI EPND	https://www.imi.europa.eu/projects-results/project-factsheets/epnd
Integration with other EBRAINS services - Dataset examples of the MIP in the EBRAINS Knowledge Graph	
CCEP database of the Medical Informatics Platform	https://search.kg.ebrains.eu/instances/Dataset/ebe50517-41d5-4029-9355-04f1e49e23c8; DOI: 10.25493/SV5Z-FSB
Example for MIP Metadata representation in the EBRAINS Knowledge Graph	https://search.kg.ebrains.eu/instances/a526f0c6-cbd4-4864-ae3a-87ad2cb7496a; DOI: 10.25493/2XWW-K4Y
Relevant links for the HIP	
Main HIP Technical Documentation Repository	https://github.com/HIP-infrastructure/hip-doc
HIP User Documentation	https://github.com/manikbh/HIP-doc
Human Intracerebral EEG Platform (HIP) Website	https://www.humanbrainproject.eu/en/medicine/human-intracerebral-eeg-platform/
Nextcloud	https://nextcloud.com/industries/healthcare/
StoryBook	https://storybook.js.org/
Holdgraf et al., Sci Data, 2019 **P2033**	https://www.nature.com/articles/s41597-019-0105-7
CCEP Data Base	https://search.kg.ebrains.eu/instances/Dataset/ebe50517-41d5-4029-9355-04f1e49e23c8
CCEP Manager	https://ccepmanager.readthedocs.io
IntrAnat for electrode localization	https://f-tract.eu/software/intranat
F-tract	https://f-tract.eu/
Health Data Cloud (HDS)	https://www.healthdatacloud.eu/

General links relevant for the EBRAINS Medical Data Analytics Service	
EBRAINS Website	https://ebrains.eu/
EBRAINS General Terms of Use	https://strapi-prod.sos-ch-dk-2.exo.io/EBRAINS_General_Terms_of_use_e457353c1a_d2122f84c2.pdf
EBRAINS Access Policy	https://strapi-prod.sos-ch-dk-2.exo.io/EBRAINS_access_policy_v1_1_ed4b84ee9e_e90146495f.pdf
EBRAINS Data Use Agreement	https://strapi-prod.sos-ch-dk-2.exo.io/EBRAINS_Data_Use_Agreement_90858e7836_ef3ee29d50.pdf
HBP Data Policy Manual	https://sos-ch-dk-2.exo.io/public-website-production/filer_public/73/ca/73ca0402-e7b9-47cc-a126-f6a193a46e7d/d1241_d781_d114_sga2_m1_accepted_190723.pdf