# Datenschutzkonforme virtuelle Forschungsinfrastrukturen für Gesundheitsdaten

Actionable Multilevel Health Data

Petra Ritter





## Human Health Data



Even after directly identifying information such as names, faces, addresses, or dates of births were removed (often called 'pseudonymization'), the data is still considered as personal data under GDPR if it could be attributed to a natural person by the use of additional information (Recital 26, GDPR) and hence the requirements for protection nevertheless apply.

Typical biomedical data like MRI or genetic material contain extensive person-related information such that re-identification cannot be excluded (Byrge & Kennedy, 2018; Gymrek et al., 2013; Rocher et al., 2019)



# Service for sensitive data: Synergistic developments



https://virtualbraincloud-2020.eu/



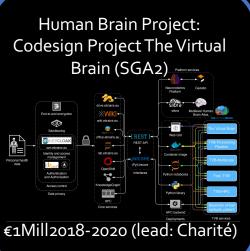
https://ebrains.eu/service/the-virtual-brain

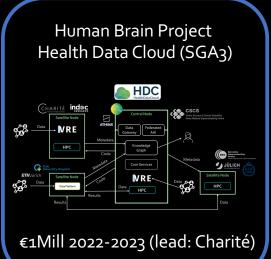


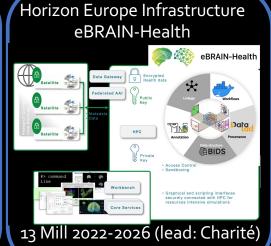
www.healthdatacloud.eu











# Service for sensitive data: Synergistic developments



https://virtualbraincloud-2020.eu/



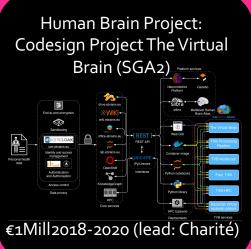
https://ebrains.eu/service/the-virtual-brain

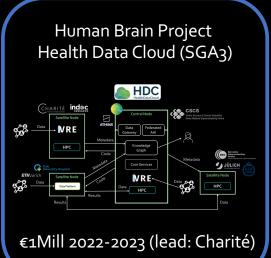


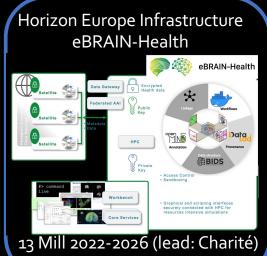
www.healthdatacloud.eu



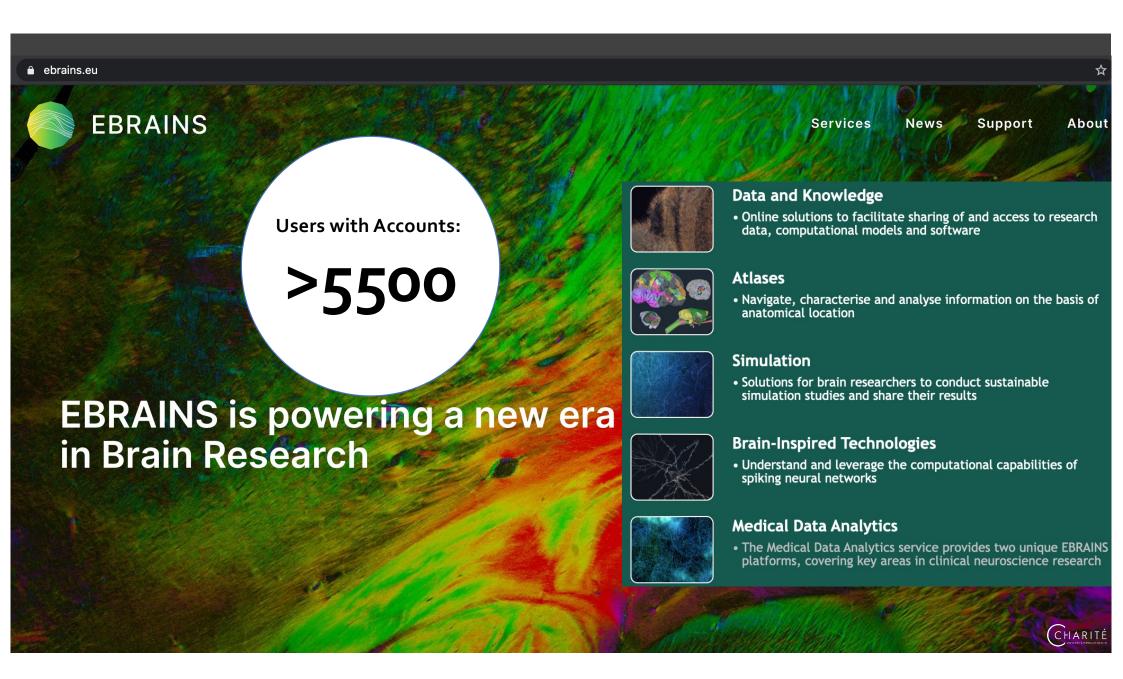




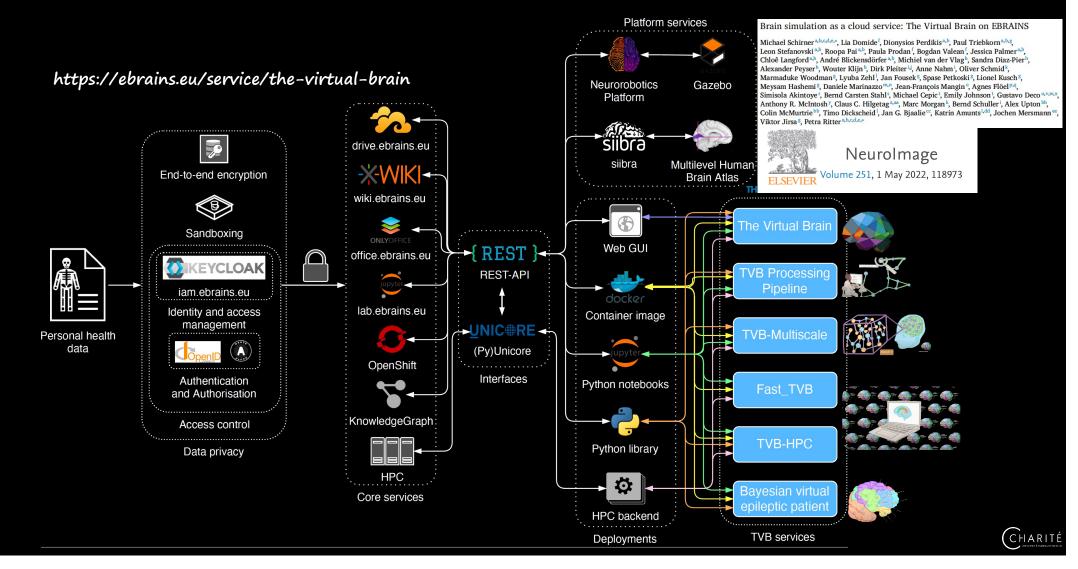




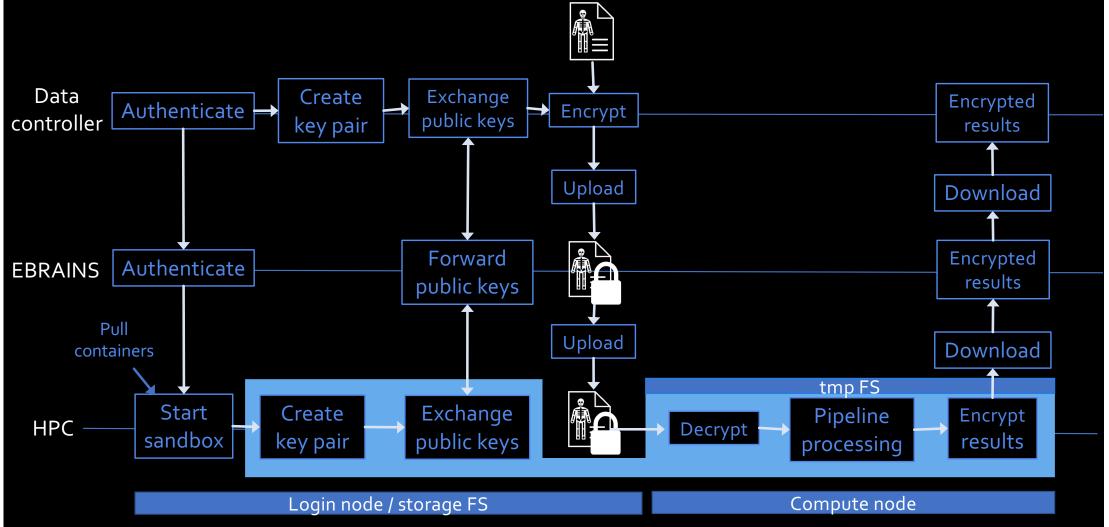




## **EBRAINS Virtual Brain Cloud**



# Automated cryptography & sandboxing workflow



**C**HARITÉ

The Virtual Brain

Patent allowance in the US (EU, Canada pending)
McIntosh, Mersmann, Jirsa, Ritter
https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2014020428



Virtual deep brain stimulation: Multiscale co-simulation of a spiking basal ganglia model and a whole-brain mean-field model with The Virtual Brain

i Jil M. Meier, Dionysios Perdikis, André Blickensdörfer, Leon Stefanovski, Qin Liu, Oliver Maith, Helge Ü. Dinkelbach, Javier Baladron, Fred H. Hamker, Petra Ritter



# Service for sensitive data: Synergistic developments



https://virtualbraincloud-2020.eu/



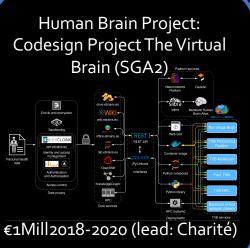
https://ebrains.eu/service/the-virtual-brain

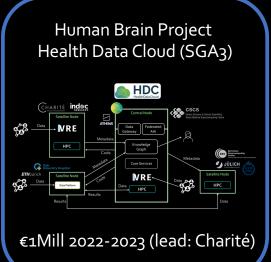


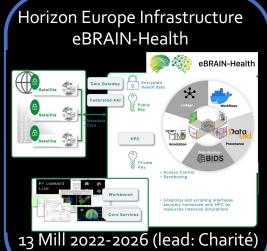
www.healthdatacloud.eu













# Service for sensitive data: Synergistic developments



https://virtualbraincloud-2020.eu/



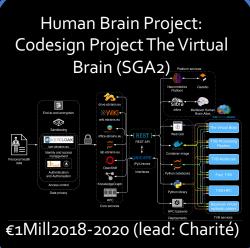
https://ebrains.eu/service/the-virtual-brain

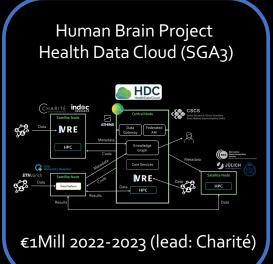


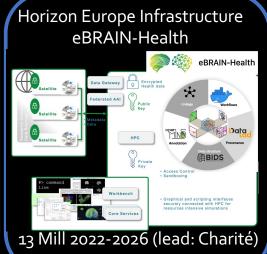
www.healthdatacloud.eu















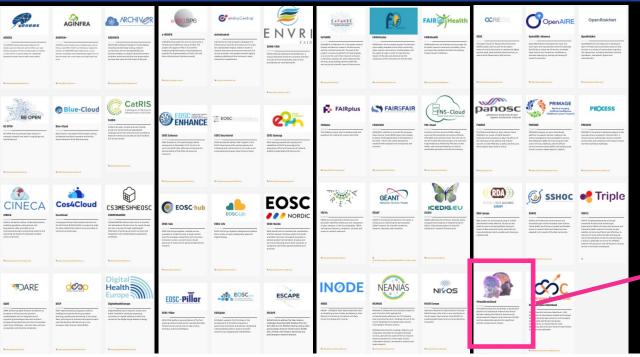
## EOSC Project Virtual Brain Cloud



bout Services & Resources Policy Use Cases Media For providers Subscribe Using the Portal 🔍

Home » About » EOSC Projects

### **EOSC Projects**



https://virtualbraincloud-2020.eu/

The Virtual Brain Cloud / €15Mill

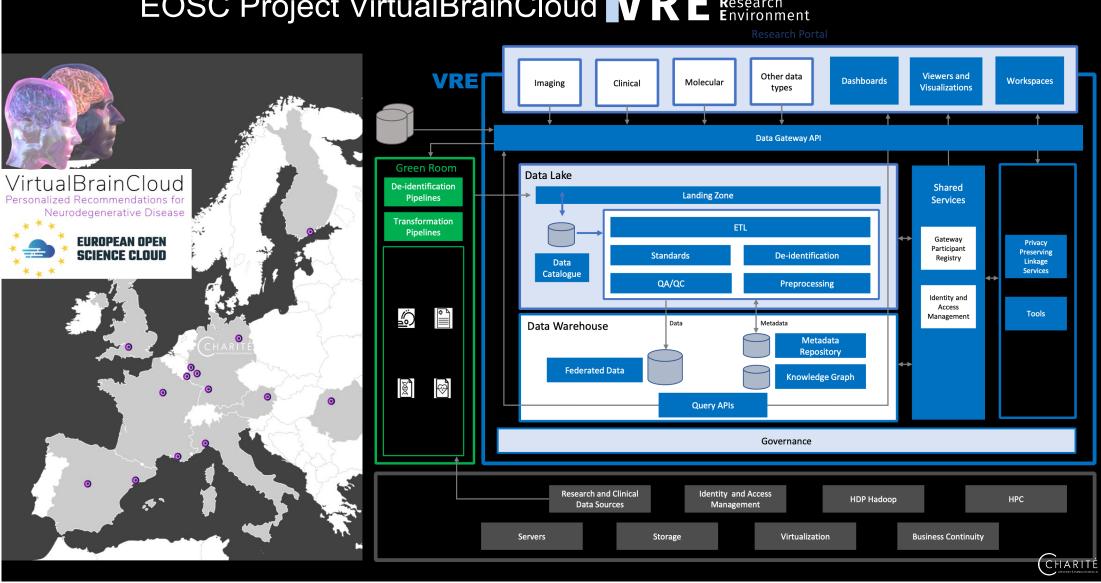
2018-2022 (lead: Charité)

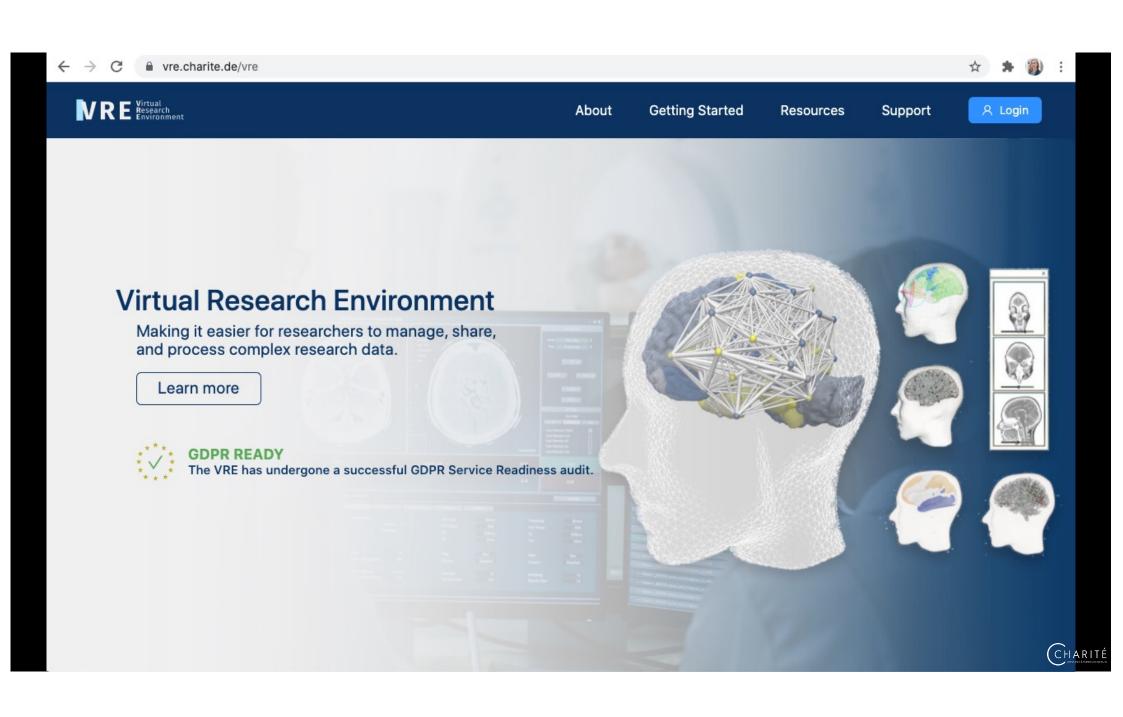
Successor: eBrain-Health / 13 Mill

2022-2026 (lead: Charité)



## EOSC Project VirtualBrainCloud VRE Research Environment





### Community

The VRE is a community development project. We follow the FAIR guiding principles for scientific data management - Findability, Accessibility, Interoperability, and Reusability.

Please contact us to learn more about how to join the VRE community and participate in the development. We want to ensure that this development is open and accessible to all clinicians and scientists at Charité, BIH, and beyond, so we can solve today's pressing medical challenges together!

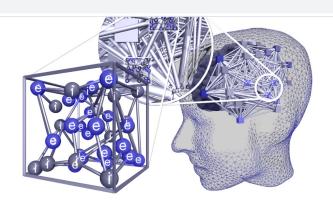
88 mbers	□ 15 Projects	品 <b>156</b> GB Data
21 Machines	<b>□ 244</b> Cores	© 1503 GB RAM



nfdi-neuro.de

neuroscience.

- o The initiative brings together data collectors, data users and technology providers to develop processes for data management in research with specific emphasis on neuroscience.
- The consortium works to build a community to develop the conceptual and practical basis of research data management for the neurosciences.



#### Latest news



#### Global survey on data-sharing barriers in neuroscience

6. October 2022 /// No Comments

The INCF Infrastructure Committee is trying to identify barriers to data sharing and reuse among neuroscience researchers worldwide, with a brief anonymous survey. The results

Read More »



#### eBRAIN-Health project awarded funding by European Union!

**BRAIN-Healtl** 

NFDI-Neuro is proud to announce that the eBRAIN-Health project: "Actionable Multilevel Health Data", coordinated by our Consortium spokesperson, Prof. Petra Ritter, has been funded

Read More »



**NFDI Neuroscience Co-Spokesperson** Thomas Wachtler contributed to the

#### **Upcoming events**

November 7 @ 09:00 - November 10 @ 19:00 CET

**EBRAINS Training Event - Simulate with EBRAINS** 

View Calendar

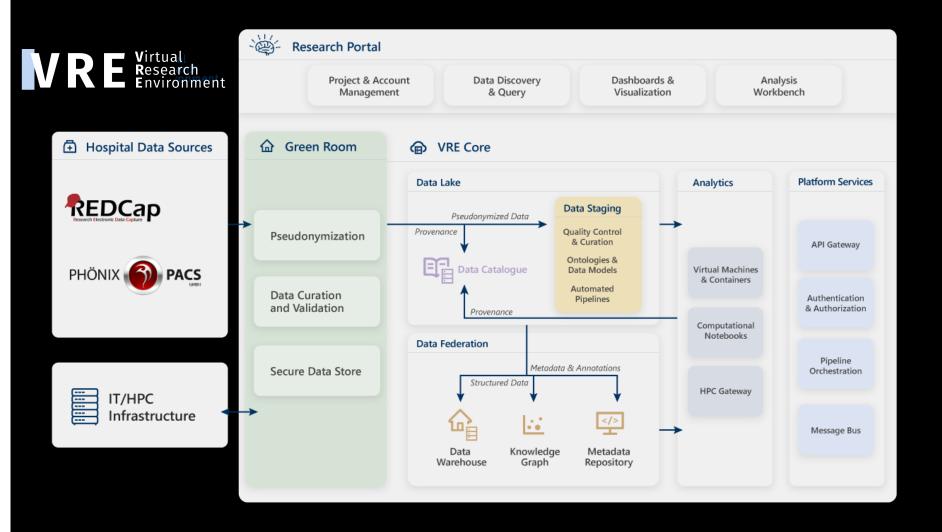


The Virtual Research Environment (VRE) is a collaborative project, developed for the research community and with the research community. VRE Community Meetings are organized for members of the research community to discuss their projects and use cases and for the development team to update on new features and development activities.

Registration link for the biweekly Virtual Research **Environment Community Meetings** 

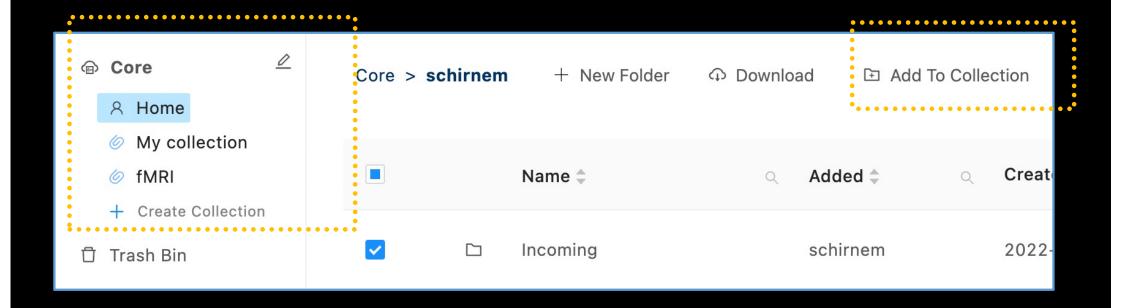


## Virtual Research Environment



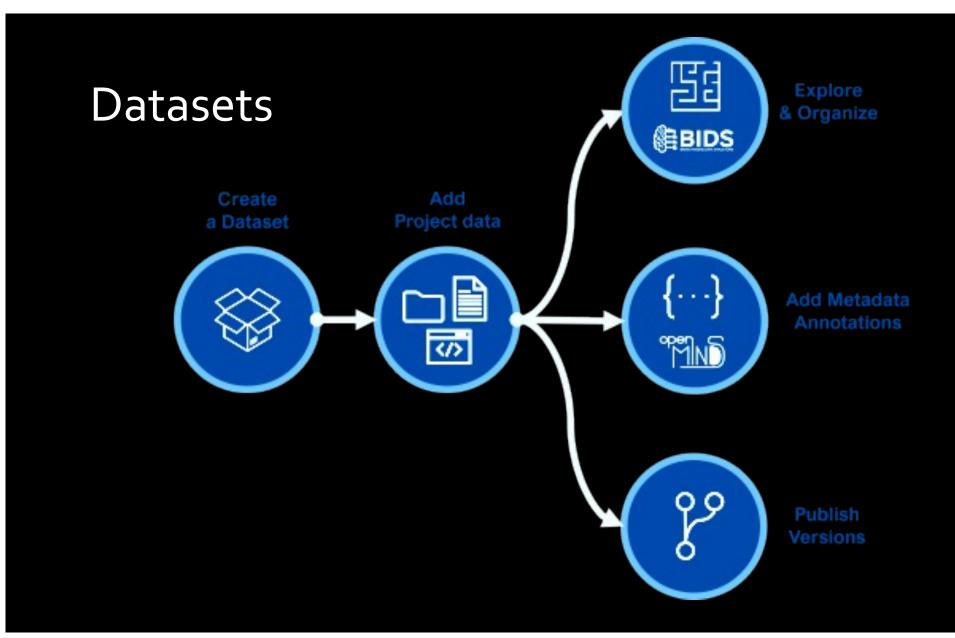


## Collections



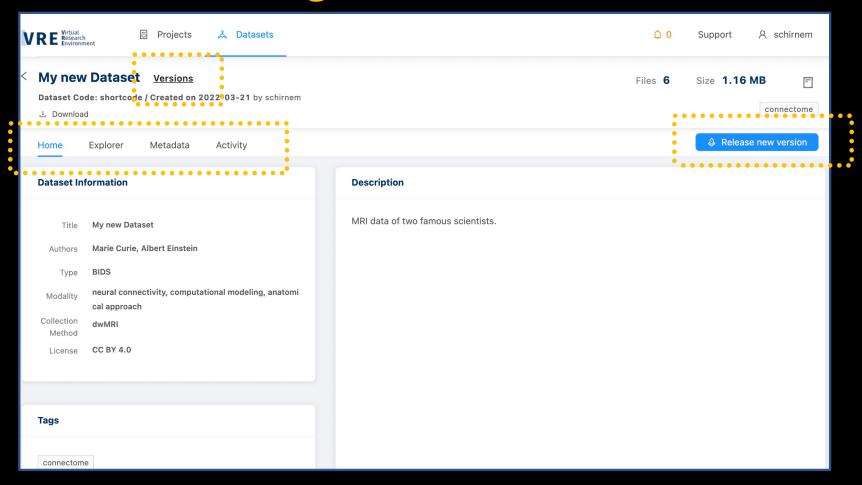
Virtual folders to group files without affecting their storage location or data lineage





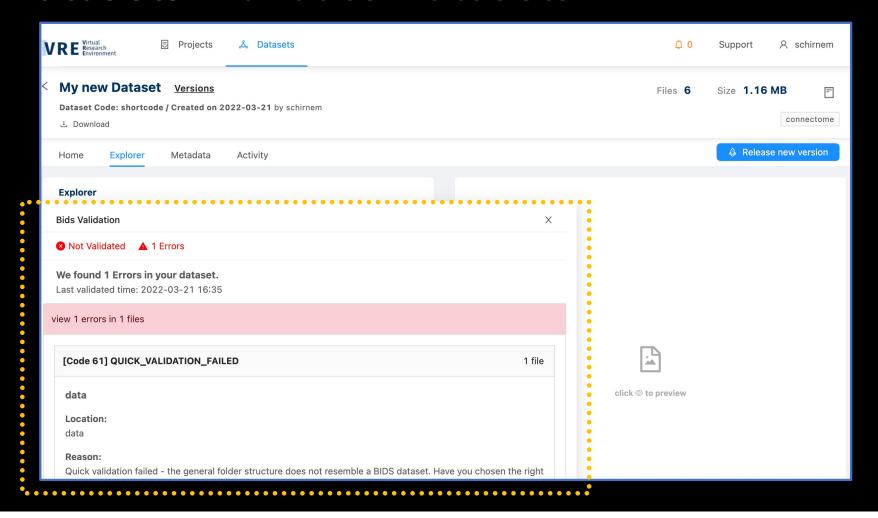


## Datasets-Manage Datasets



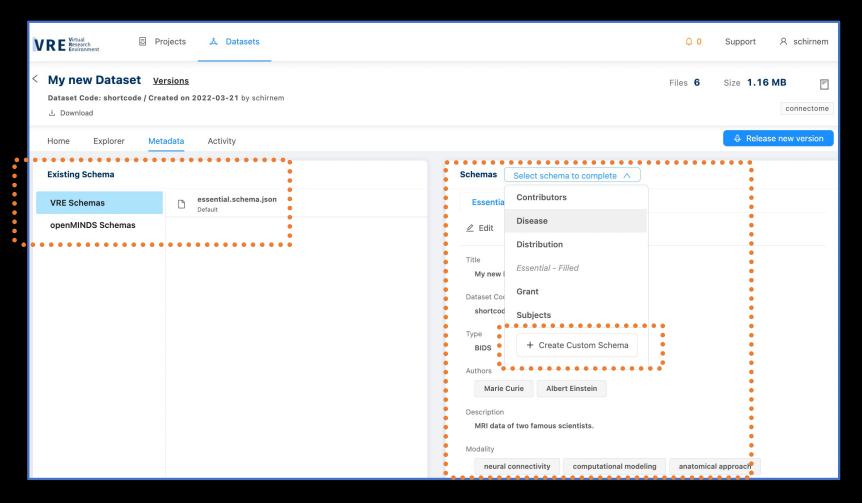


## Datasets-Validate Datasets



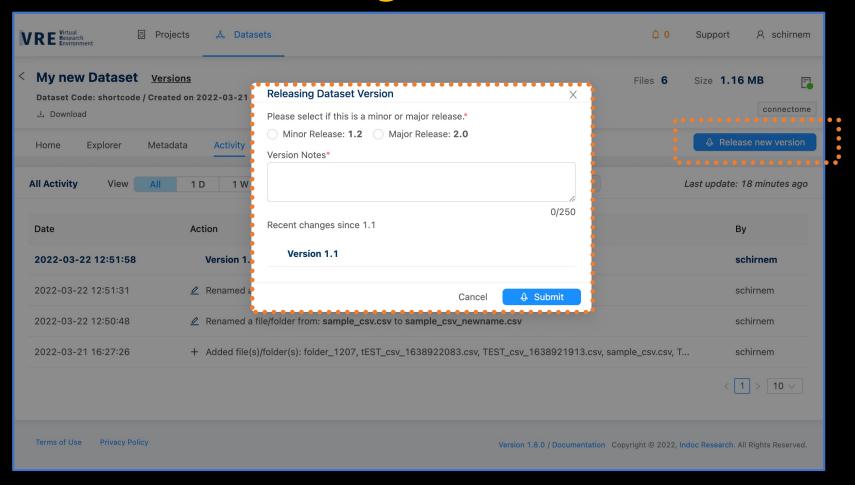


## Datasets-Metadata schemas





## Datasets-Versioning





## Data lineage





#### **Current File**

Represents the current state of the file in its lineage. Hover over the file node to view the *Type, Name,* and *Process Time* of the current file.



### **Processing Node**

The Processing node indicates a processing activity such as a processing pipeline, or copy/delete action. Hover over the Processing node with your cursor to view the processing action and the date and time it was completed.



#### **Upstream or Downstream Version of Current File**

Represents either an upstream or downstream version of the current file. Hover over the file node to view the *Type, Name,* and *Upload Time* of the file.



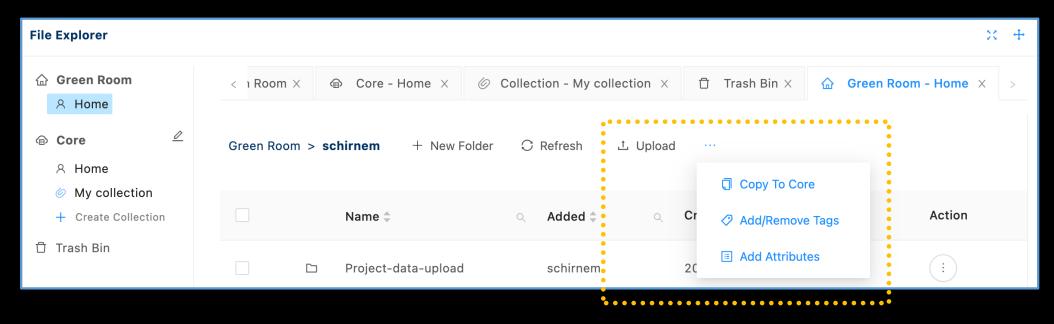
#### **Data Transfer Node**

This node represents data transfer or copy from the Green Room to the Core.





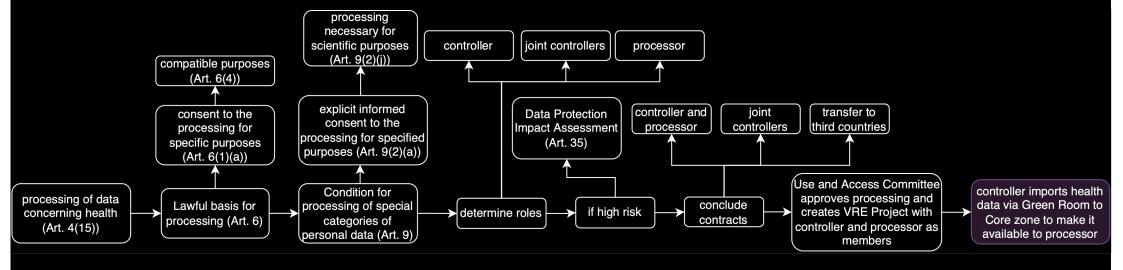
# Data onboarding—User roles



- Project Contributors can only upload files to isolated (user-specific) spaces in the Green Room or private Core zone
- Only Project Administrators can make data accessible to the entire Project team
- All files that can be accessed by a user, can also be downloaded by that user



## T&O measures for lawful processing

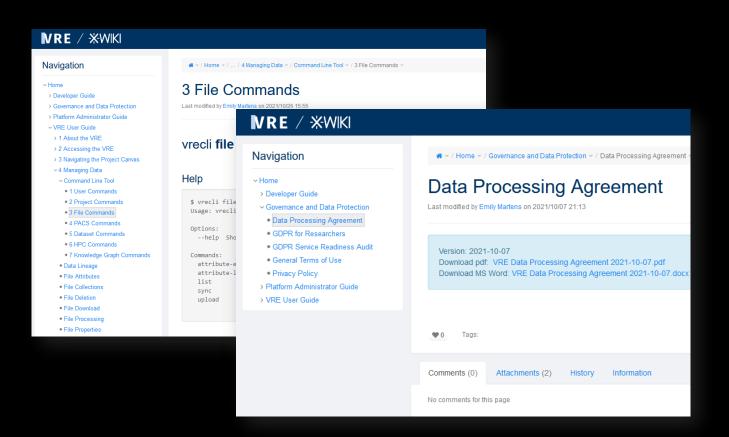


- Only after the Use and Access Committee approved all necessary contracts a dedicated VRE Project is created by a Platform Administrator and adds the specified controller and processor as team members
- Controller can then import the purpose-specific health data to make it available to processor for performing the processing operation



## Wiki with support resources

- User guide
- Policies and templates to assist researchers with GDPR compliance





### T&O measures—IT infrastructure

- VRE is hosted at the IT center of the Berlin Institute of Health at Charité
- As a designated 'critical infrastructure' of the German government Charité IT implements state of the art security measures and industry best practices (certification every two years)
- independent data protection evaluation of the VRE completed in July 2021 concluded that the services of the VRE can be offered as commissioned data processing for health-related research projects in compliance with the requirements of the European GDPR



# T&O measures—Isolation of data and resources

- VRE restricts direct access to its resources such as storage and compute
- Users can interact with resources only via controlled routes via the GUI and the CLI
- Kubernetes namespaces to create isolated zones
- Additionally, zones are hosted on different VMs; traffic restricted by network filters
- Copy To Core can only be authorized by Project Administrators and uses a multi-step protocol to prevent accidental exposure



## T&O measures—Encryption

- Data are encrypted during transmission to/from the VRE
- Data stored ("at rest") on the VRE are encrypted using disk-level (hardware) encryption



# T&O measures—Authentication and authorization

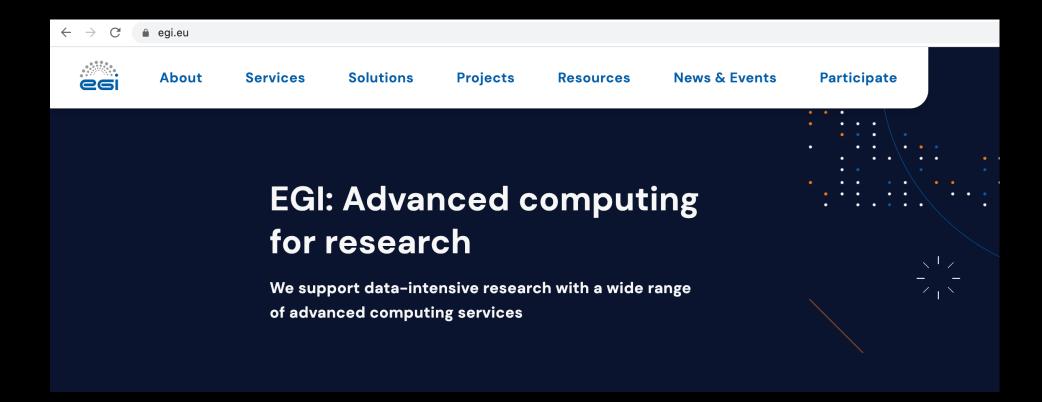
- Keycloak/OpenID Connect for single sign-on and to authenticate communications between the VRE front-end, API Gateway (which connects all back-end services), and workbench tools
- Identity of VRE users federated between the VRE identity and access management system (IAM) Keycloak and the Charité IAM Microsoft Active Directory (AD)
  - Registration in the Charité AD is a precondition for using the VRE
  - external users must complete a registration procedure and provide a form of personal identification and evidence of a contractual relationship with the Charité to be authorized for inclusion in the Charité AD
  - Account information and credentials are stored in the Charité AD and not within the VRE itself
- Password complexity requirements and session inactivity timeouts reduce the risk of unauthorized access



### T&O measures—Access control

- Fine-grained role-based and project-based access control governed by transparent Use and Access Policies that promote fair and democratic access to the VRE
- Onboarding a new project follows a streamlined procedure: Service Agreement, Data Processing Agreements, Terms of Use, and positive Research Ethics Board evaluation
- After approval, Project Administrator is first granted access by invitation from the Platform Administrator; Project Administrator can then invite specified users to access the project data
- a VRE user may be a member of multiple projects and may have different roles in each project











# Service for sensitive data: Synergistic developments



https://virtualbraincloud-2020.eu/

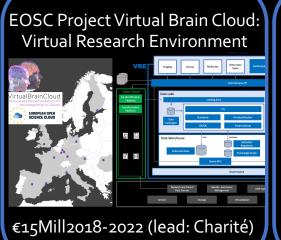


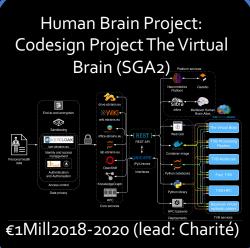
https://ebrains.eu/service/the-virtual-brain

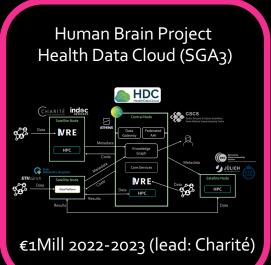


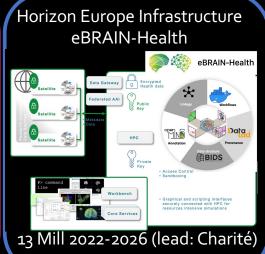
www.healthdatacloud.eu







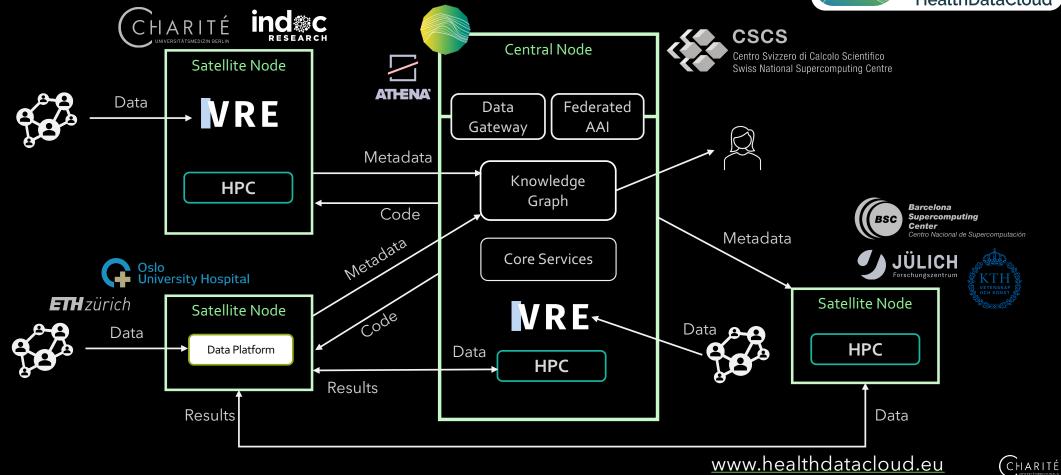






## EBRAINS Health Data Cloud





## Governance

#### Supervisory **Authorities**

EU, Federal, State

#### **Processor**

Maintains the Data Platform.

Implements Technical and Organizational Measures

Processes data on behalf of the controller.

### **Data Protection**

and processor



#### **Data Platform**

A secure data processing environment

### Officer

Designated by the controller

Involved in issues relating to protection of personal data.



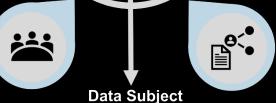
#### Controller

Determines the purpose and means of processing

Prepares a data protection impact assessment for the research study

#### Research Ethics Board

Ethical evaluation of the research project



### Dataset use terms

Data Processing Agreement with receiving data controller

**Derivative Datasets** 

An identified or identifiable natural person whose personal data is the subject of processing.

Consents to their data being used and shared for research



# Service for sensitive data: Synergistic developments



https://virtualbraincloud-2020.eu/

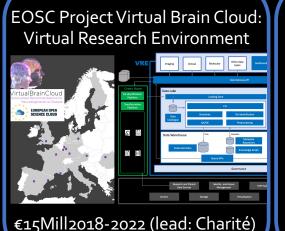


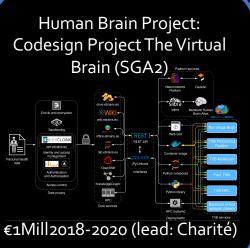
https://ebrains.eu/service/the-virtual-brain

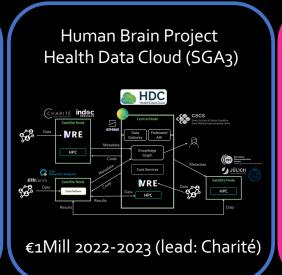


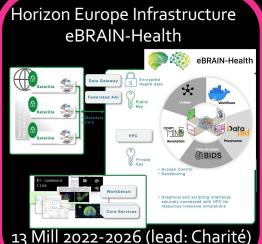
www.healthdatacloud.eu





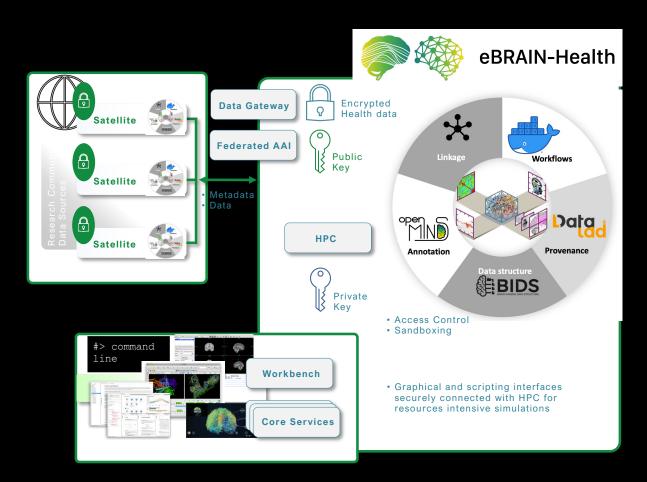








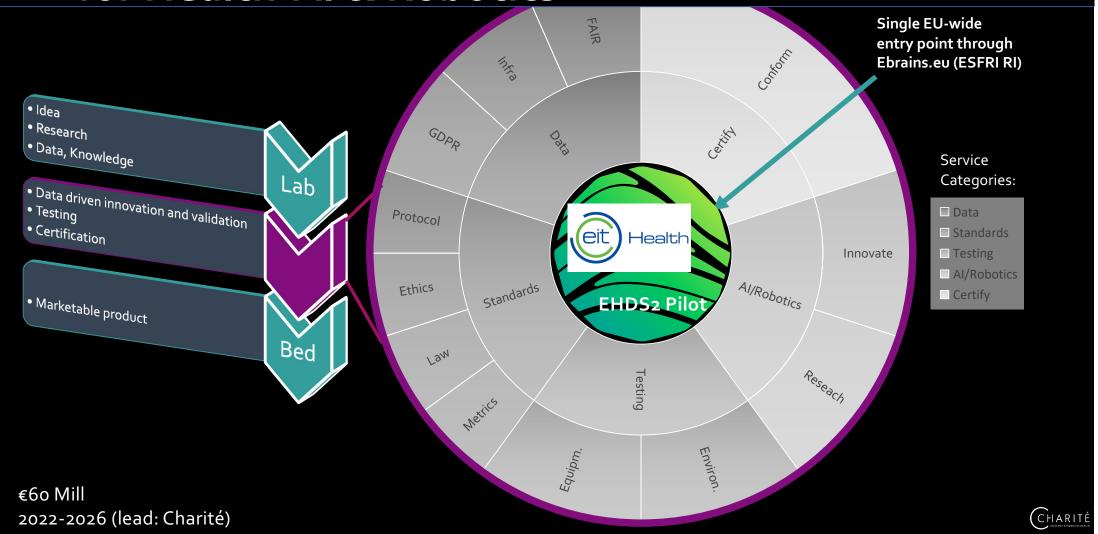
## EBRAIN-Health



In collaboration with:



# Action: Closing the Gap in the Innovation Chain for Health-AI & Robotics



## World Health Summit 2022



OCTOBER 16-18 BERLIN, GERMANY & DIGITAL











### PD 11 - How to Achieve a Global Health Data Space

Oct. 18, 2022, CEST: 09:00 AM - 10:30 AM / UTC: 07:00 AM - 08:30 AM

Panel Discussion (PD 11) - Europe

Breaking the data silos is essential to reshape the future of healthcare and crisis preparedness. digital health and artificial intelligence (AI), we are, more than ever before, in the pole position that could help to treat and govern data for health as a global public good. However, across an divide in the capacity to effectively work with data. The 2021 I-DAIR Global Research Map reve leaders based in a small number of countries and the rest of the world is growing. Additionally, race, and age limit the universal benefit and the trust in data use.

Only when used equitably and ethically, the work with data can offer an unprecedented possible well-being and achieve UHC 2030. Impact at the global level will thereby only be possible with from different disciplines including from countries of relatively lower income and/or with smalle we wish to spark stakeholder engagement for a transparent data system that protects citizens' healthcare and research.

#### Chairs:

Prof. Dr. Petra Ritter

Charité - Universitätsmedizin Berlin | Berlin Institute of Health (BIH) | Professor for Brain Simu

#### Speakers:

Dr. Marlies Dorlöchter

DLR Project Management Agency | International Health Research | Head of Division | Germany

<u>Dr. Mehdi Snène</u>

International Digital Health and Artificial Intelligence Research Collaborative (I-DAIR) | CEO ad

Paweł Świeboda

Human Brain Project | Director-General | Belgium



