

The HBP Collaboratory

a solution to facilitate collaboration in neuroscience

Marc MORGAN, Swiss Federal Institute of Technology in Lausanne (EPFL)
HBP Project Coordination Office – Technical Coordination – Geneva

Collaboration enablers

1. Identifying people you want to collaborate with
2. Locating a controlled workspace for exchange
3. Documenting goals, ideas, processes, progress
4. Sharing data
5. Being able to try out different ideas interactively
6. Knowing that what was done is reproducible
7. Sharing results and workflows

The Collaboratory is evolving

- During the first years of HBP, we developed **Collaboratory 1.0**
 - Today it counts several thousand users
- Since the beginning of HBP, **cloud technology** has evolved significantly
- We are currently rolling out **Collaboratory 2.0**
 - Based on open source tools with a wide base of users/developers
 - More user friendly, more features
 - Adapts to policies of the EBRAINS research infrastructure produced by HBP

Identifying people you want to collaborate with

- Collaboratory provides a [Single Sign-On \(SSO\)](#) solution to access HBP services
 - One account, one username, one password
- Based on industry standard OpenID Connect (OIDC) protocol
 - Collaboratory 2.0 uses [KeyCloak](#), an open source tool, supported by RedHat/IBM
- Extending to Single Sign On over a very large network of institutions with [eduGain](#) (provided by GEANT, also known for eduroam)
- Serving a [community](#) of users. We are adding a [directory](#) of users.

A **controlled** workspace for exchange

- The Collaboratory is **self-hosted** by EBRAINS
- We target deployment at the FENIX Research Infrastructure sites:
 - **CSCS, Swiss National Supercomputing Center (Lugano, Switzerland)** ← Today
 - BSC, Barcelona Supercomputing Center (Barcelona, Spain)
 - CEA, Alternative Energies and Atomic Energy Commission (Paris, France)
 - CINECA (Bologna, Italy)
 - JSC, Jülich Supercomputing Center (Jülich , Germany)

A controlled **workspace** for exchange

- Any group of users can create a workspace = a **collab**
- A collab is either completely **public** (crawlable by Google) or **private** to a group of users
- **Permissions** are defined for each collab individually
 - Admin, edit, and read permissions
 - For individual users and/or other groups

Documenting goals, ideas, processes, progress

The Collaboratory encourages 2 modes of documentation:

- **Wiki** pages to create web pages for simple access
- **Collaborative edition** of **Office** documents: Word, Excel, Powerpoint
 - Compatible with MS-Office formats
 - Collaborative edition *a la* Google Docs provided by OnlyOffice
- Any other format is also supported in the storage of the collab.

Sharing data

- Each collab comes with its own **storage** space...
- ... but data can be **shared** across collabs too (given the user has the permissions)
- All the data is **version controlled** for easy access to the change history
- EBRAINS policies will define storage **quotas**.
- Note: EBRAINS policies forbid uploading non-**anonymized human** data

Interactive and reproducible workflows

- Based on [Jupyter Notebooks](#), a Python programming environment
- With access to a very large set of Python scientific [libraries](#)
- Users can try out code snippets [interactively](#)
- Collaboratory 2.0 builds on [JupyterLab](#) with support for more complex workflows
- Notebooks can access [storage](#) & [computing](#) resources at the supercomputing sites.

Extending the Collaboratory

- The Collaboratory is an **extensible** framework
- Developers can add **Community apps** to offer specialized services
- Apps are supported by Collaboratory **services**:
 - identification, drive, collabs, permissions
- Community apps can be **integrated** inside wiki pages

Sharing results and workflows

- Open science, open data, open source ...
- ... but you choose when to open access to others

- The Collaboratory is a great solution for:
 - Collaborative research in a remote team setting
 - Education: can be used for MOOCs
 - Live papers: give readers of your publication direct access to the data you used and let them reproduce your workflows

Collaboratory: under the hood

- User identification & authentication (IAM): [KeyCloak](#)
- Wiki pages: [XWiki](#)
- File storage (Drive): [SeaFile](#)
- Collaborative edition of Office documents: [OnlyOffice](#)
- Reproducible workflows: [JupyterLab](#)

- Runs on [OpenStack](#) (SWIFT) & [Kubernetes](#)

Thank you for your attention

Questions: marc.morgan@epfl.ch
or: support@humanbrainproject.eu