

LE DATALAB EN COUPE

Des utilisateurs au socle d'infrastructure : une coupe technique du Datalab, strate par strate.

Mode de lecture : comme une coupe géologique, chaque bande est une couche technique, lue de haut en bas — de la surface (utilisateurs) jusqu'au socle physique (clusters). Les flèches indiquent qu'une couche s'appuie sur la suivante.

UTILISATEURS

- 🔍

Chercheurs
- 📊

Data scientists
- 👥

Équipes projet

Chercheurs, data scientists et équipes projet accèdent au Datalab pour explorer, analyser et visualiser les données de l'EDS.

SE CONNECTE À

LE DATALAB – APPLICATIONS & RESSOURCES DE CALCUL

OUTILS

- 📄

JupyterLab
notebooks interactifs (code, texte, résultats)
- 🖼️

OHIF
viewer web d'imagerie médicale (DICOM)
- ⚙️

SLURM
ordonnanceur de jobs batch sur le cluster

EXÉCUTE DU CODE VIA

LANGAGES & MOTEURS

- <>

Python
langage généraliste, data science
- ∧

R
analyse statistique
- ⚡

Spark
calcul distribué en mémoire
- 🗄️

HiveQL
requêtes SQL sur Hadoop

LIT / ÉCRIT SUR

STOCKAGE

- 💾

Bloc
disque virtuel attaché à une VM
- 📁

Objet
fichiers volumineux, accès via API
- 📄

Fichiers (FS)
système de fichiers partagé entre nœuds

S'EXÉCUTE SUR LE SOCLE

CLUSTERS

- 🐘

Hadoop
calcul & stockage distribués (HDFS, Spark, Hive)
- 🏠

Kubernetes
orchestration des conteneurs — héberge JupyterLab, OHIF...
- 🔄

CEPH
stockage distribué unifié : bloc, objet, fichiers
- 🖨️

Cluster GPU
nœuds de calcul accélérés par GPU

Chaque usage du Datalab s'appuie sur un type de stockage, lui-même porté par l'un des clusters d'infrastructure ci-dessus.