

DOMAINE : Bioinformatique

INTITULE : Pratiques FAIR en bio-informatique

Objectifs généraux de formation

Introduction aux bonnes pratiques en bio-informatique afin de pérenniser son travail de recherche.
Notion de reproductibilité, répétabilité du code. Bonnes pratiques de partage et gestion des versions des outils utilisés.
Cette formation permet de découvrir les bonnes pratiques dans le cadre d'un travail nécessitant des approches programmatiques (statistiques, programmation d'outils, analyses de données biologiques).
Elle s'inscrit dans l'aspect science ouverte afin de rendre plus facilement disponible le travail bio-informatique.

Contenu

- Introduction aux pratiques FAIR
- Archivage et gestion de code avec Git
- Encapsulation de son code et gestion de l'environnement avec conda, singularity et snakemake
- Documentation de son code avec Rmarkdown et Jupyter

Programme

- **J1 / J2**
 - Introduction générale aux principes FAIR (Faisabilité, Accessibilité, Interopérabilité et Reproductibilité)
 - Introduction à la gestion de code via Git
 - TP pratique sur Git
- **J3 / J4**
 - Introduction à la notion d'encapsulation et de gestion de l'environnement avec conda, docker et singularity
 - TP sur la gestion des environnement et des paquets avec conda
 - TP sur la gestion des images et des conteneurs avec singularity
 - TP sur la gestion des workflow avec snakemake
- **J5**
 - Introduction générale à la documentation de son code avec la programmation lettrée
 - TP rédaction d'un rapport d'analyse avec Rmarkdown
 - TP utilisation de l'environnement Jupyter Lab

Modalités Pédagogiques

Formation avec une phase de présentations théoriques et une phase de travaux pratiques pour appliquer les notions présentées dans la formation.

Avoir un compte sur le cluster de calcul du Mésocentre Clermont-Auvergne (faire une demande le cas échéant sur le site <https://hub.mesocentre.uca.fr>)

VENIR AVEC UN ORDINATEUR PORTABLE muni d'une connexion à Eduroam opérationnelle.

Public concerné

Cette formation s'adresse à un public de bioinformaticiens ou de biologistes ayant des notions de programmation ou d'utilisation des lignes de commandes et qui cherchent à améliorer leurs pratiques.

Durée et rythme

5 journées de formation
9h- 12h, 14h-17h

Calendrier

lundi 3 avril 2023
mardi 4 avril 2023
mardi 11 avril 2023
mercredi 12 avril 2023
jeudi 13 avril 2023

Formateur

Pierre MARIN
Nadia GOUÉ
Matéo HIRIART
Philippe RUIZ

Lieu

UCA - Campus des Cézeaux - Bâtiment Turing - Mésocentre - 7 rue Blaise Pascal - Salle A009