

Jeudi 13 novembre - Visio



Présentation de la formation

Thomas Denecker <https://orcid.org/0000-0003-4253-1048>

Laurent Jourden <https://orcid.org/0000-0003-4253-1048>

Claire Toffano-Nioche <https://orcid.org/0000-0003-4134-6844>





MERIT (Métier en bioinformaTique) est le **réseau national** des ingénieurs en bioinformatique des établissements d'**Enseignement Supérieur et de la Recherche** (ESR).

MERIT est soutenu par plusieurs instituts du **CNRS**: Sciences informatiques, Biologie, Écologie & Environnement, Chimie, Terre & Univers.

Ce réseau est **ouvert à l'ensemble des personnels**, de toutes les tutelles, intéressés par les animations métier en ingénierie bioinformatique (ingénieurs, chercheurs,...).





Le réseau **MERIT** et l'**Institut Français de Bioinformatique** (IFB) organisent une Action Nationale de Formation (ANF) sur les principes FAIR appliqués aux langages de workflows bioinformatiques à destination des bioinformaticien·ne·s



Cette formation est portée par **INS2I** (CNRS Sciences Informatiques), avec un soutien financier de l'**INRAE** et la participation de **Hélène Chiapello** et **Vincent Lefort** du bureau MERIT





À la fin de cette formation, vous :

- Aurez acquis les bases et connaîtrez certaines fonctions avancées des gestionnaires de workflows
- Aurez écrit un workflow (proposé par l'équipe pédagogique) en Nextflow et connaîtrez les bases pour le transposer en Snakemake
- Aurez exécuté le workflow sur le cluster mis à disposition par l'IFB et l'aurez optimisé par IA





- Disposer d'un **ordinateur portable** pendant toute la formation
- Avoir une bonne maîtrise d'**Unix/shell** : gestion de fichiers et répertoires, lancement de commandes, scripting
- Avoir une bonne maîtrise de **Git** en ligne de commande (pull, add, commit, push, log, branch, switch, checkout...)
- Utiliser une plateforme de partage de code comme **GitHub, GitLab** et savoir créer une branche, évaluer et valider une pull request.
- Maîtrise d'au moins un outil de workflow (**Nextflow** ou **Snakemake**)
- Avoir suivi le tutoriel Nextflow : <https://training.nextflow.io/latest/> ou pratiquer régulièrement Nextflow



- Avoir un compte sur le cluster de l'IFB

<https://my.cluster.france-bioinformatique.fr/manager2/login>

- Avoir un compte sur gitlab.com





Intro Générale



Redéfinition d'un workflow



Présentation de l'exemple appliqué "fil rouge" de la formation



Pause



Introduction à la reproductibilité au sens large



Pause déjeuner



Quelles différences entre Conda/Pixi, les containers, Spack et Guix ?



Création d'un environnement "conda/Pixi" et packaging d'un outil



Pause



Enrichissement d'une image d'un container avec cet environnement



Pause



Pourquoi utiliser un gestionnaire de workflow, Comparaison Snakemake / Nextflow



Galaxy: souvent oublié et pourtant !



 Introduction Nextflow (Seqera / nf-core / Nextflow)

 Données de référence

 Utilisation données distantes

 Pause

 Flux conditionnel

 Exploiter la puissance HPC

 Pause déjeuner



Création d'un module, tests



Création d'un module, dépôt sur nf-core



Pause



l'IA avec Nextflow



Passage du dev à la prod



Concepts workflow avancés : reporting



Traduction des développements Nextflow avancés en Snakemake



Enrichir le code en métadonnées pour le rendre plus FAIR



Pause



Comment les entrepôts permettent d'améliorer la FAIRness des workflows



Partage : Archiver son code dans Software Heritage, le diffuser dans HAL



Pause déjeuner



Rendre plus FAIR son dépôt Gitlab



Pause



Conclusion



Retour à chaud

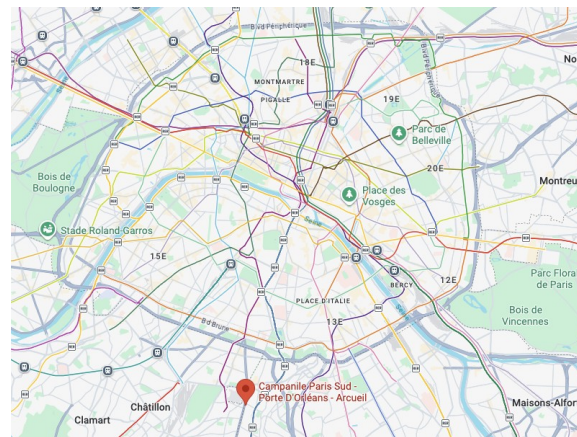
La formation aura lieu en résidentielle au Campanile Paris Sud (formation, hébergement et repas).

Adresse

73 Avenue Aristide Briand
94110 Arcueil, France

Transport

RER B, Ligne 197





Pour les chambres

Plus informations à venir de la part de l'hôtel

Bagagerie disponible



Repas

- Petit déjeuner (pour ceux qui logent) de 06:30 - 09:30
- Déjeuner : Buffet + café
- Diner : Buffet + Boisson





Les cours : <https://moodle.france-bioinformatique.fr/course/view.php?id=39>

Les codes : <https://gitlab.com/anf-workflow-et-reproductibilite>

Des questions ?

