

Science ouverte et principe FAIR dans un projet de bioinformatique, 28-30 juin 2021

<https://ifb-elixirfr.github.io/IFB-FAIR-bioinfo-training/session2021.html>

Formateurs : G. Le Corguillé, J. Seiler, T. Denecker, C. Hernandez, C. Toffano-Nioche

Helpers : Hugo Arbes, Emilie Drouineau, Pauline François et Paulette Lieby

Organisatrices : H. Chiapello, Y. Mahmah

L'IFB est l'infrastructure nationale de bioinformatique qui assure un support, déploie des services, organise des formations et réalise des développements innovants pour les communautés des sciences de la vie et de la santé.

Missions

- **Développer un service national**
 - **Infrastructure de services en bioinformatique**
 - **Développement et déploiement de ressources logicielles (outils, bases de données)**
 - **Formations** en bioinformatique, pour les biologistes et les bioinformaticiens
 - **Support aux programmes de recherche** en biologie, santé, agronomie, environnement
 - Accès aux **technologies de pointe**
 - Mutualisation des **compétences et expertises** de haut niveau
- **Développer une vision stratégique** pour maintenir la recherche française au plus haut niveau
 - Axes d'innovation: **bioinformatique intégrative, intelligence artificielle**
 - **Science ouverte**: accessibilité, interopérabilité, gestion des données, reproductibilité des résultats, ...
 - Servir de levier pour concevoir des projets ambitieux aux niveaux national et international
- **Représentation internationale de la bioinformatique française** (notamment auprès d'ELIXIR)

Compute and storage

- National Network of Computing Resources (**NNCR**)
- High-performance computing (**HPC**)
- Dual environment: **cloud + cluster**
- Data protection policy (**GDPR**)

2020: 11 PB storage, 21K cores

Software development and deployment

- **Development** of specialised software
- **Packaging** (conda)
- **Virtualisation** (appliances, containers)
- **Good practices** for software engineering
- **Open and collaborative** code

2020: >900 Tools

Consulting and support to researchers

- From the conception to the deliverables
- Data management plans (**DMP**)
- Design and implementation of **workflows**
- **Data science**

A centralised helpdesk

- A **unique entry point** for all IFB services and platforms.
- A network of **400 experts** covering **all fields of bioinformatics** (sequence, structure, networks, biostatistics, ...).
- Coverage of **all research domains**: fundamental biology, agriculture, health, environment, biotechnologies



Training

- **Thematic schools** (NGS & other omics,, biostat, programming, workflows)
- **Bring Your Own Data (BYOD)** training courses
- FAIR bioinfo and FAIR data training
- **Diploma in integrative bioinformatics**
- Adaptation to a fast-evolving demand

Integrative bioinformatics

- **Multi-omics + structure + imaging + ...**
- **Statistical approaches**
- **Multiplex networks**

"Comment rendre un projet bioinformatique d'analyse et développement plus reproductible"

Lundi 29 juin 2021 : 10:00-18:00

- 10:00 - 10:30 **Présentation de la formation**
- 10:30 - 11:30 **Introduction à la Reproductibilité**
- 11:30 - 13:00 **Sauvegarde du code (git, Github) : cours & TP**
- 13:00 - 14:00 *Déjeuner*
- 14:00 - 14h30 **TP Github**
- 14:30 - 15:30 **Contrôler son environnement de développement (Conda, Docker)**
- 15:30-16:00 **TP Conda**
- 16:00 - 16:30 *Pause*
- 16:30- 18:00 **Usage d'une image Docker : cours & TP**

Mardi 30 juin 2021 : 9:00-17:30

- 9:00 - 10:30 **Gestion de pipelines (snakemake) : cours & TP**
- 10:30 - 11:00 *Pause*
- 11:00 - 13:00 **Traçabilité et partage (notebook, markdown): cours & TP**
- 13:00 - 14:00 *Déjeuner*
- 14:00 - 15:00 **Les usages sur un cluster HPC (IFB Core Cluster) : Conda, Singularity**
- 15:00 - 15:30 **Déploiement d'un workflow sur l'IFB Core Cluster (TP)**
- 15:30 - 16:00 *Pause*
- 16:00 - 17:00 **Exposer son projet**
- 17:00 - 17:30 *Conclusion*

Mercredi 30 juin 2021 : 9:00-16:00

Atelier optionnel de travail sur le matériel pédagogique de la formation à destination de futurs formateurs

- 9:00 - 10:00 Bilan des évaluations des participants. Points forts et points à améliorer
- 10:00 - 12:00 Travail sur le matériel de cours (diapos, exercices)
- 12:00 - 13:00 *Déjeuner*
- 13:00 - 15:00 Travail sur le matériel de cours (diapos, exercices)
- 15:00 - 16:00 Restitution et Conclusion

Merci de remplir le questionnaire **avant la fin des 2 jours de formation** :

<https://forms.gle/2Tjxa9kejKmksS9C7>

Librement inspiré de : <https://openlifesci.org/code-of-conduct>

- **Soyons amical et patient.** Nous sommes ici entre collègues. Demain vous serez (peut-être) formateur
- **Soyons accueillant.** Pas de discrimination d'aucune sorte ici !
- **Soyons prévenant.** Respectez le travail des autres
- **Soyons respectueux.** Restons calme et courtois en cas de désaccord
- **Essayons de comprendre pourquoi nous ne sommes pas d'accord.** Essayons de résoudre les points de désaccord et divergence de manière constructive. Par ailleurs l'erreur est humaine... et reproductible (en général)

L'équipe pédagogique

Les participants

Vous ? Votre équipe / unité ?

Votre première expérience de la reproductibilité et des principes FAIR ?

Vos attentes ?