

13 au 15 juin 2022, Institut des Systèmes Complexes, Paris

FAIR Bioinfo 2022

Les principes FAIR dans un projet de bioinformatique

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

Helpers : H. Arbes, E. Drouineau, P. François

Organisatrices : H. Chiapello, Y. Mahmah





Objectifs

- Rendre une analyse bioinformatique reproductible
- Introduire et pratiquer les principes “FAIR” et outils associés
- Enrichir le matériel pédagogique et former des formateurs

Bref historique

- 1ère édition locale à l’I2BC en 2019 (C. Toffano-Nioche et T. Denecker)
- Partenariat avec l’IFB et mise en pratique sur l’infra IFB
 - [Edition 2020](#) du 31 août au 2 sept à l’ISC Paris
 - [Edition 2021](#) du 28 au 30 juin à l’ISC Paris
 - Edition 2022 du 13 au 15 juin à l’ISC Paris

Publication : <https://jose.theoj.org/papers/10.21105/jose.00068>



LE NNCR : National Network of Computing Ressources

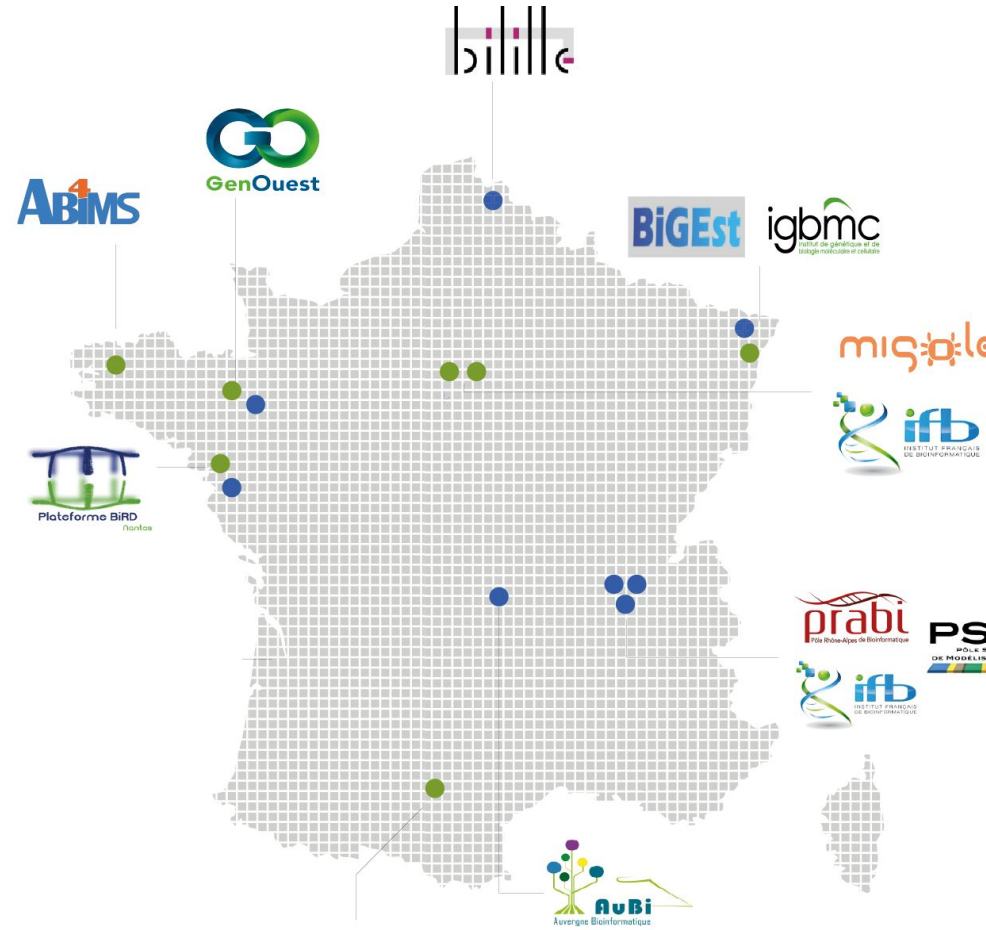
2 infrastructures complémentaires : **Cloud** et **Cluster**

Réseau national composé :

- D'équipements de plateformes régionales IFB
- De 2 infrastructures Core (IFB core cluster + IFB core cloud)

Administré par une "task force" mutualisée

- Les plateformes partagent X% d'ETP d'ingénieurs
- Bénéfice mutuel des % d'ETP partagés
 - Forte émulation du personnel
 - Éviter la déperdition d'énergie
 - Partager l'expérience
 - Partager les développements
 - Traitement collaboratif des problèmes
 - Solutions robustes



cluster

- 6 SLURM Clusters
- 16,900 vCPU
- 92 TB RAM
- 10.5 PB
- 10 GPU cards
- 4935 yearCPU in 2020

8 sites clouds avec 8,000 vCPU et 36To RAM

● Cluster
● Cloud



Catalogue de nos formations : <https://www.france-bioinformatique.fr/formations-agenda/>

FAIR bioinfo

Objectifs

- > Fondamentaux de la **reproductibilité** en bioinformatique
- > Introduire et pratiquer les outils qui permettent d'adopter les principes "FAIR"
- > Mise en pratique sur l'infra IFB

Modalités

- > **Présentiel** à Paris, 100 €/j académique, 400 €/j sinon
- > 2 jours + 1 (optionnel pour futurs formateurs)
- > 3 éditions 2020, 2021, 2022

FAIR data

Objectifs

- > Fondamentaux de l'**Open Data** (pratiques, juridiques)
- > Le PGD dans un contexte bioinformatique
- > Les standards de métadonnées pour les « omiques »

Modalités

- > **Distanciel** synchrone, gratuit
- > 4 ou 5 demi-journées
- > 2 éditions distancielles nationales en 2021, 2 éditions présentiels régionales en 2022



Céline Hernandez, I2BC



Claire Toffano-Nioche, I2BC



Thomas Denecker, IFB



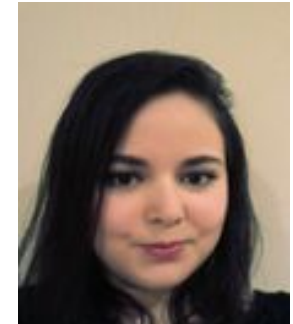
Gildas Le Corguillé, IFB



Julien Seiler, IFB



Hélène Chiapello, IFB



Yousra Mahmah, IFB



Emilie Drouineau, I2BC



Hugo Arbes, I2BC



Pauline François, CIRI



Tout est sur moodle [ici](#)

▪ **Lundi 13 juin - 9:30-17:30**

Accueil, présentation de la formation, tour de table

Introduction FAIR & reproductibilité

Usecase 1 - e-labbook

Organiser son projet, Notebook, concept des liens entre texte et code (JupyterLab local - Démo live)

Versionning de code (Git)

Usecase 2 - Développement (en local, exposition)

Premiers pas vers un environnement contrôlé : Conda. Exposer son travail (GitHub/GitLab). Intégration github et jupyter

DOI avec Zenodo. GitHub Pages

▪ **Mardi 14 Juin - 9:30-17:00**

Distribuer son code et l'environnement (Docker)

Usecase 3 - cluster

Cluster de calcul de l'IFB (Jupyterlab, slurm). Workflow (Snakemake). Singularity

Conclusion - Questionnaires évaluation

▪ **Mercredi 15 juin - 9:30-16:00 (journée optionnelle pour futur formateurs)**

Bilan des évaluations des participants et discussions sur les points d'amélioration

Travail sur le matériel de cours (diapos, exercices)

Travail sur le matériel de cours (diapos, exercices)

Conclusion

Déjeuner sur place avec plateaux repas : 12:45-13:45



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

- **Soyons amicaux et patients.** Nous sommes ici entre collègues. Demain vous serez (peut-être) formateur
 - **Soyons accueillant.** Pas de discrimination d'aucune sorte ici
 - **Soyons prévenants.** 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)
- + [Toutes les questions sont bienvenues !](#)



Les participants

Vous ? Votre équipe / unité ?

Votre première expérience de la reproductibilité et des principes FAIR ?

Vos attentes ?