

# Luiz Felipe Piochi

[in](#) luiz-piochi [✉](mailto:luizpiochi@outlook.com) luizpiochi@outlook.com [G](#) lupiochi [L](#) lupiochi.me

## FORMATION

**UNIVERSITÉ DE LORRAINE - IAEM** | *Doctorat en Informatique* 2024 – présent

- Thèse : "Apprentissage de l'interactome de surface hôte-pathogène pour la conception de nouvelles thérapies."

**UNIVERSITÉ DE COIMBRA - DCV** | *Master en Biologie Cellulaire et Moléculaire* 2020 – 2023

- Moyenne : 19/20; Cours approfondis en Bio-informatique, Statistiques; Classé dans le top 3% de l'Université

## EXPÉRIENCE PROFESSIONNELLE

**INRIA - CAPSID** Nancy, France  
**Docteurant** Septembre 2024 – présent

- Utilisation de l'apprentissage profond avancé pour la conception de protéines et le développement de médicaments
- Élaboration d'un modèle d'apprentissage profond pour la prédiction des interactions protéine-ligand

**ERBEL.EUS** Bilbao, Espagne  
**Data Scientist en Génomique** Juin 2024 – Septembre 2024

- Participation à un projet européen et un projet national sur l'apiculture et la métagénomique
- Création de deux pipelines automatisés de métagénomique pour le traitement des données NGS, accélérant le flux de travail interne de plus de 70
- Traitement et analyse de plus d'1 To de données de séquençage métagénomique

**BIOCANT - LABORATOIRE DE CONCEPTION MOLÉCULAIRE BASÉE SUR LES DONNÉES** Cantanhede, Portugal  
**Chercheur en Bio-informatique** Novembre 2021 – Décembre 2023

- Développement et publication d'un outil d'apprentissage profond pour la prédiction de la résistance aux médicaments dans une revue de bio-informatique à fort impact
- Traitement et analyse de données multi-omiques issues de plus de 10 jeux de données publiques et de centaines de lignées cellulaires
- Mise en place d'un pipeline pour le criblage virtuel de plus de 300 composés contre des protéines d'intérêt dans le neurodéveloppement

**CNC-UC - MITOLAB** Coimbra, Portugal  
**Chercheur Étudiant** Octobre 2019 – Juillet 2020

- Réalisation de multiples essais métaboliques et caractérisation de la dynamique de différentes lignées cellulaires cancéreuses
- Développement d'un pipeline Python interne pour l'analyse et la visualisation de données, réduisant de 30
- Compilation et publication du projet sous forme de revue dans un journal de référence sur le vieillissement

## COMPÉTENCES

**Programmation** Python, R, Bash, SQL, JavaScript, Snake-make **Langues** Portugais (langue natale), Anglais (courant), Espagnol (C2), Français (A0), Allemand (A1)

**Frameworks** numpy, pandas, torch, tensorflow

## PROJETS

**GORGON** 2024

- Python [Numpy, Pandas], R [Phyloseq, ggplot2], Snakemake
- Développement d'un pipeline de métagénomique modulaire et évolutif pour le traitement et l'analyse rapides de grands jeux de données NovaSeq
- Optimisation des processus complexes, du prétraitement des données brutes à la production de rapports détaillés, générant des résultats rapides
- Conçu pour assurer la reproductibilité dans différents environnements informatiques et la gestion efficace de jeux de données de tailles variables

**DELFO** 2023

- Python [Numpy, Pandas, TensorFlow], R [Tidyverse, Seurat]
- Développement d'un modèle d'IA apprenant à partir de la chimie des médicaments et de données multi-omiques, y compris des données unicellulaires, pour prédire la réponse cellulaire à différents composés
- Performances supérieures aux algorithmes de pointe de l'époque, avec des métriques de régression élevées à travers plusieurs étapes de validation
- Publication dans Oxford University Press Bioinformatics