



## **Offre de Stage - Sciences des Données en Oncologie Création d'une Base de Données en Oncologie : Intégration Multimodale et Application de Techniques de Machine Learning**

*Durée : 6 mois (avril 2025 - septembre 2025)*

*Localisation : Institut Paoli-Calmettes (IPC) en collaboration avec SESSTIM, Aix-Marseille  
Université (AMU)*

### **Contexte et objectifs :**

Ce stage s'inscrit dans un projet de recherche visant à intégrer des données cliniques, d'imagerie, et génomiques pour construire une base de données avancée en oncologie. Le but est de faciliter l'analyse des données et d'améliorer les prédictions sur la survie et la récurrence du cancer du sein.

Le stagiaire sera impliqué dans :

1. **Collecte et nettoyage des données** : Extraction et harmonisation de données issues de multiples sources.
2. **Constitution de la base de données** : Conception et structuration d'un entrepôt de données robuste pour des analyses futures.
3. **Application de techniques d'intelligence artificielle** : Détection d'anomalies, fusion de données, et analyse prédictive avec des algorithmes de machine learning.

### **Environnement du stage :**

Le stage se déroulera au sein du nouvel entrepôt de données de santé de l'IPC, offrant un accès à des outils et technologies avancés en science des données et en intelligence artificielle. Le stagiaire travaillera sous la supervision conjointe de :

- **Dr. Louis Tassy**, oncologue à l'IPC.
- **Dr. Raquel Urena**, maître de conférences spécialisée en IA pour la santé au SESSTIM.
- **Dr Matie Guyomard**, post doctorante en IA et santé.

### **Conditions de stage :**

- **Indemnité de stage** : Rémunération statutaire selon les barèmes d'Aix-Marseille Université.
- **Durée** : 6 mois à temps plein.
- **Possibilité d'évolution** : À l'issue du stage, le/la stagiaire pourrait se voir offrir l'opportunité de participer à un projet de recherche plus large, avec une possibilité de financement pour une thèse en lien avec le sujet.

### **Profil recherché :**

- Étudiant(e) en Master 2 en Sciences des Données, Informatique ou équivalent.
- Connaissances solides en Python et en bases de données (SQL, NoSQL).
- Intérêt pour le traitement des données en santé et la bio-informatique.
- Esprit analytique et capacité à travailler en équipe.

### **Compétences développées :**

- Expertise en intégration et préparation des données multimodales en santé.

- Maîtrise des processus ETL (Extraction, Transformation, Chargement).
- Application des techniques de machine learning en oncologie.

**Candidature :**

Envoyez CV et lettre de motivation à [raquel.urena@univ-amu.fr](mailto:raquel.urena@univ-amu.fr), [marie.guyomard@univ-amu.fr](mailto:marie.guyomard@univ-amu.fr), [tassyl@ipc.unicancer.fr](mailto:tassyl@ipc.unicancer.fr) avant le **15 février 2025**.