



**Offre n°2025-09057**

## **Doctorant F/H Modèle de fondation tabulaire**

**Type de contrat :** Fixed-term contract

**Niveau de diplôme exigé :** Graduate degree or equivalent

**Fonction :** PhD Position

**Niveau d'expérience souhaité :** Recently graduated

### **A propos du centre ou de la direction fonctionnelle**

Le centre de recherche Inria de Saclay a été créé en 2008. Sa dynamique s'inscrit dans le développement du plateau de Saclay, en partenariat étroit d'une part avec le pôle de l'**Université Paris-Saclay** et d'autre part avec le pôle de l'**Institut Polytechnique de Paris**. Afin de construire une politique de site ambitieuse, le centre Inria de Saclay a signé en 2021 des accords stratégiques avec ces deux partenaires territoriaux privilégiés.

Le centre compte **40 équipes-projets**, dont 32 sont communes avec l'Université Paris-Saclay ou l'Institut Polytechnique de Paris. Son action mobilise **plus de 600 personnes**, scientifiques et personnels d'appui à la recherche et à l'innovation, issues de 54 nationalités.

Le centre Inria Saclay - Île-de-France est un acteur essentiel de la recherche en sciences du numérique sur le plateau de Saclay. Il porte les valeurs et les projets qui font l'originalité d'Inria dans le paysage de la recherche : l'excellence scientifique, le transfert technologique, les partenariats pluridisciplinaires avec des établissements aux compétences complémentaires aux nôtres, afin de maximiser l'impact scientifique, économique et sociétal d'Inria.

### **Contexte et atouts du poste**

Cette proposition de thèse se concentre sur les modèles de fondation tabulaires : des modèles d'apprentissage automatique pré-entraînés pour les données tabulaires. En particulier, cette proposition se concentrera sur des architectures neuronales bien adaptées aux données relationnelles d'entrée, des données avec des types multiples (chaînes, nombres, dates...) et souvent réparties sur plusieurs tables. Pour modéliser ces données, nous utiliserons une architecture flexible basée sur les réseaux attentionnels graphiques. Le défi sera de faire le bon choix de graphe et d'architectures correspondantes afin d'appliquer le contexte pertinent.

## Mission confiée

### Missions :

Avec l'aide de Gael Varoquaux, la personne recrutée fera de la recherche doctorale en apprentissage statistique pour développer des modèles tabulaires.

The goal of this PhD proposal is to build learning architectures that can fit very diverse types of a relational data, in order to 1) be pretrained on massive sources of data, 2) transfer to new tables and relational databases without requiring too much data wrangling. For this, the starting point will be to build a transformer-based architecture that models naturally the various input types: strings, dates, and numbers. Further work will adapt this architecture to multiple tables.

## Principales activités

Principales activités :

- Faire de la recherche sur les modèles tabulaires, et en particulier les modèles de fondation
- Suivre la littérature dans le domaine
- Publier les résultats et les présenter

## Compétences

Solide bagage mathématique: statistiques, machine learning

Maîtrise de Python.

La connaissance de PyTorch est un plus.

Esprit curieux.

Bonne maîtrise écrite de l'anglais

## Avantages

- Restauration subventionnée
- Transports publics remboursés partiellement
- Congés: 7 semaines de congés annuels + 10 jours de RTT (base temps plein) + possibilité d'autorisations d'absence exceptionnelle (ex : enfants malades, déménagement)
- Possibilité de télétravail et aménagement du temps de travail
- Équipements professionnels à disposition (visioconférence, prêts de matériels informatiques, etc.)
- Prestations sociales, culturelles et sportives (Association de gestion des œuvres sociales d'Inria)
- Accès à la formation professionnelle
- Sécurité sociale

## Rémunération

2200€ brut/mois

## Informations générales

- **Thème/Domaine** : Optimization, machine learning and statistical methods Statistics (Big data) (BAP E)
- **Ville** : Palaiseau
- **Centre Inria** : [Centre Inria de Saclay](#)
- **Date de prise de fonction souhaitée** : 2025-09-01
- **Durée de contrat** : 3 years
- **Date limite pour postuler** : 2025-08-31

## Contacts

- **Équipe Inria** : [SODA](#)
- **Directeur de thèse** :  
Varoquaux Gael / [Gael.Varoquaux@inria.fr](mailto:Gael.Varoquaux@inria.fr)

## A propos d'Inria

Inria est l'institut national de recherche dédié aux sciences et technologies du numérique. Il emploie 2600 personnes. Ses 215 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3900 scientifiques pour relever les défis du numérique, souvent à l'interface d'autres disciplines. L'institut fait appel à de nombreux talents dans plus d'une quarantaine de métiers différents. 900 personnels d'appui à la recherche et à l'innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 200 start-up. L'institut s'efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

**Attention:** Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.

## Consignes pour postuler

### Sécurité défense :

Ce poste est susceptible d'être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L'autorisation d'accès à une zone est délivrée par le chef d'établissement, après avis ministériel favorable, tel que défini dans l'arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un poste affecté dans une ZRR aurait pour conséquence l'annulation du recrutement.

### Politique de recrutement :

Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.