

## Data Scientist / Ingénieur Données

Le médialab de Sciences Po recrute un ingénieur de recherche pour travailler sur l'analyse de l'écosystème médiatique français

### Contexte

Le poste s'intègre dans le projet GOPI (Geometry of Public Issues), projet ANR dont l'objectif est de modéliser la dynamique des conversations en ligne dans l'espace public numérique.

Notre équipe collecte déjà massivement les traces d'activité sur Twitter et Facebook et ce à travers différents pays. Nous utilisons des outils d'analyse de réseau pour saisir les ancrages idéologiques dans lesquels se déploient les débats publics en ligne. En outre, nous suivons quotidiennement des centaines de médias français pour suivre l'évolution des principaux cadres d'interprétation qui structurent différentes controverses socio-techniques.

Le projet GOPI s'appuie sur une équipe hautement interdisciplinaire, comprenant des politologues, des sociologues, des designers, des mathématiciens, des linguistes (PNL) et des ingénieurs. Nous nous appuyons sur des recherches de pointe en matière d'analyse de données pour faire progresser la recherche en sciences sociales et politiques, en mettant l'accent sur les phénomènes en ligne tels que les mouvements sociaux, la polarisation, l'extrémisme, les chambres d'écho et les bulles de filtre.

### Principaux objectifs du poste

L'objectif principal de ce poste est de développer des outils logiciels, de mener des opérations de collecte dans une grande variété de plateformes numériques et de participer à des analyses de données. L'ingénieur de recherche qui sera engagé participera à deux types d'activités :

La collecte de données : Nous avons pour objectif d'améliorer la collecte de l'activité des réseaux sociaux en France et dans le monde, de développer les outils et l'architecture permettant d'accroître notre activité.

Développement de la science des données : Nous visons à renforcer la capacité de notre équipe à implémenter des algorithmes en science des données dans les méthodes d'analyse de réseau et d'apprentissage automatique en 1) refaisant le code existant en outils

performants, 2) conceptualisant de nouveaux algorithmes, et 3) implémentant les algorithmes proposés dans les articles scientifiques.

Vous pouvez consulter des exemples de nos logiciels de collecte et d'analyse de données sur notre page Github.

## Compétences et expérience souhaitées

- Python est le langage de programmation privilégié dans l'équipe, mais coder dans un autre langage n'est pas rédhibitoire
- Rédaction du code dans des environnements collaboratifs.
- Expérience de travail avec la collecte de données en ligne (outils de collecte, API).
- Expérience de travail sur des serveurs distants et manipulation de grands jeux de données.
- (si possible) Expérience de la mise en œuvre de codes d'analyse de données provenant de la littérature de recherche scientifique et des conférences de data science.
- (si possible) Expérience de l'écriture de modules Python.
- (si possible) Expérience des données et de l'analyse des réseaux sociaux.
- (si possible) Expérience antérieure avec TensorFlow et Pytorch bienvenue.
- (si possible) Une expérience avec les bases de données (SQL) et avec les outils pour grands volumes de données (recherche élastique, stockage distribué, mise en file d'attente) serait très appréciée.
- (si possible) Expérience ou formation en sciences des données. L'intérêt pour la formation continue (cours en ligne) est apprécié.
- (si possible) Expérience en matière de mise en réseau des données scientifiques.
- (si possible) Un doctorat en informatique est préférable, mais d'autres formations équivalentes sont les bienvenues. Les bons profils de master et d'ingénieur sont également invités à candidater.

## L'environnement de travail

Vous rejoindrez l'équipe technique du médialab SciencesPo à Paris et travaillerez avec des chercheurs, des ingénieurs, des concepteurs et des spécialistes des données. C'est un bon environnement de travail pour apprendre un large éventail d'outils de science des données dans les domaines de l'apprentissage automatique, de l'analyse de réseau, de textes avec des applications en sociologie et en sciences politiques.

Les ingénieurs de recherche sont encouragés à participer en tant que co-auteurs à notre production d'articles scientifiques et à mener leurs propres projets de publication s'ils le souhaitent.

Salaire compétitif en fonction de l'expérience.

Il s'agit d'un poste de 12 mois, débutant en janvier 2021 ou plus tôt en fonction des disponibilités. Possibilité de renouvellement.

Nous nous adaptons à l'évolution des conditions de travail imposées par la pandémie de COVID19, le travail à distance à temps partiel est possible.

## Pour postuler

Envoyez un CV et une lettre de motivation détaillant votre motivation et vos compétences pour le poste et le cas échéant, un lien vers vos publications scientifiques récentes (publiées ou en préparation) / projet de développement sur github avant le 30 novembre 2020 à cette adresse électronique :

[recrutement.medialab@sciencespo.fr](mailto:recrutement.medialab@sciencespo.fr).

Pour toute demande de renseignements sur le poste :

[pedro.ramaciottimorales@sciencespo.fr](mailto:pedro.ramaciottimorales@sciencespo.fr)

[jean-philippe.cointet@sciencespo.fr](mailto:jean-philippe.cointet@sciencespo.fr)

Le candidat peut commencer dès janvier 2021. Toutefois, il existe une certaine flexibilité. Veuillez indiquer dans votre candidature les dates provisoires auxquelles vous seriez disponible pour commencer.