

Référence de l'annonce : DATA-CDD-2025-06

Poste ouvert aux personnes en situation de handicap

POSTE

Poste	Data Scientist EU4HEALTH (f/h)
Catégorie d'emploi	1
Emploi repère	Chargé de projet et d'expertise scientifique en santé publique (f/h)
Type de contrat	CDD de 12 mois de droit public à plein complet. Détachement fonction publique d'Etat, Territoriale ou Hospitalière possible.
Motif	Création – poste sur convention de projet
Rémunération	Selon l'expérience et le niveau de formation par référence aux grilles indiciaires des agences sanitaires, en application du décret n°2003-224 du 7 mars 2003, ou selon statut particulier si fonctionnaire.*
Prise de fonction	Dès que possible
Renseignements	Sur le poste : Matthieu Hanf – Directeur de la Direction DATA, matthieu.hanf@santepubliquefrance.fr Sur le recrutement : Marie-Christine Morel, Chargée du recrutement à la DRH – Tél : 01 41 79 68 28

LOCALISATION GEOGRAPHIQUE

Adresse	Le siège de l'agence est situé à Saint-Maurice (94). L'agence dispose d'implantations régionales auprès des agences régionales de la santé. Localisation du poste : siège de l'agence. Site de Saint-Maurice (siège) : 12 rue du Val d'Osne à Saint-Maurice (94).
----------------	--

PRESENTATION DE L'AGENCE

Santé publique France est l'agence nationale de santé publique française. Etablissement public de l'Etat sous tutelle du ministre chargé de la santé, créée par l'ordonnance 2016-246 du 15 avril 2016, elle intervient au service de la santé des populations. Agence scientifique et d'expertise du champ sanitaire, elle a pour missions :

- 1° L'observation épidémiologique et la surveillance de l'état de santé des populations ;
- 2° La veille sur les risques sanitaires menaçant les populations ;
- 3° La promotion de la santé et la réduction des risques pour la santé ;
- 4° Le développement de la prévention et de l'éducation pour la santé ;
- 5° La préparation et la réponse aux menaces, alertes et crises sanitaires ;
- 6° Le lancement de l'alerte sanitaire.

L'Agence est organisée autour de quatre conseils (Conseil d'Administration, Conseil scientifique, Comité d'Ethique et de Déontologie et Comité d'orientation et de dialogue), de directions scientifiques et transversales, et de directions assurant le support et le soutien à l'activité. L'Agence dispose d'implantations régionales (Cellules régionales) auprès des agences régionales de la santé.

Son programme de travail, arrêté par son Conseil d'administration, s'articule autour de 6 enjeux : Anticipation, préparation et réponse aux menaces de santé publique, dont les épidémies ; Numérique en santé publique ; Santé environnementale, changement climatique et environnement de travail ; Fardeau des maladies et de leurs déterminants, efficacité des interventions et retour sur investissement de la prévention ; Stratégie de prévention, marketing social et approche par populations ; Inégalités sociales, vulnérabilités territoriales.

AFFECTATION

Direction	Direction Appui, traitements et analyses des données (DATA)
	Le poste s'inscrit au sein de la Direction Appui, Traitements et Analyses des données (DATA) de Santé publique France, qui assure un appui transversal à l'ensemble de l'agence pour la collecte, le traitement, l'analyse et la valorisation des données. Forte d'une cinquantaine d'agents, la direction est structurée en trois unités spécialisées (unité « Applications, big data et surveillance syndromique », unité « Appui et méthodes pour les études et investigations dans le domaine de la surveillance », unité « Enquêtes »). Elle intervient sur l'ensemble du cycle de vie des données de santé.
	Ses missions couvrent notamment la collecte et la gestion de données, l'analyse statistique, la géomatique, la métrologie, ainsi que le développement d'outils informatiques d'analyse et de visualisation. Elle pilote ou soutient plusieurs dispositifs structurants, notamment le système de surveillance syndromique SurSaUD, l'enquête

Baromètre santé, le site open data Odissé ainsi que l'exploitation de bases médico-administratives comme le SNDS.

La direction DATA développe une expertise reconnue en modélisation spatio-temporelle, détection automatique de signaux, et intelligence artificielle. Soucieuse de renforcer la qualité scientifique de ses travaux, elle collabore activement avec des partenaires institutionnels et académiques, et accueille régulièrement stagiaires, internes, doctorants et chercheurs, dans une dynamique d'innovation continue au service de la surveillance en santé publique.

L'une des activités de la direction est d'identifier et exploiter de nouvelles sources de données permettant à l'agence de répondre à ses missions de surveillance.

L'épidémie de Covid-19 a mis en lumière un défaut de collecte et d'exploitation de données en France et dans d'autres pays européens, notamment des données hospitalières, pour monitorer en temps proche du réel le fardeau d'une ou plusieurs pathologies pesant sur le système de soins. Si les données sont bien saisies et stockées par les différents systèmes d'information hospitaliers, celles-ci restent peu exploitées à des fins de santé publique. Sans une amélioration importante des infrastructures de surveillance préexistantes, les États membres de l'Union européenne, dont la France, risquent d'être confrontés à des défis similaires à ceux observés au début de la crise Covid-19.

Afin d'y remédier, la Commission européenne a proposé de financer, de manière conséquente, le renforcement des capacités de surveillance épidémiologique des états membres, à hauteur de 97 M d'Euros via une subvention directe (Direct Grant) issue du programme de travail 2023 de l'instrument financier « EU4Health ».

Santé publique France a obtenu un financement conséquent pour renforcer son système de surveillance épidémiologique à l'hôpital en exploitant des données stockées au sein des entrepôts de données de santé hospitaliers. Le projet visant à renforcer la surveillance à l'hôpital a été baptisé « Orchidée » (Organisation d'un Réseau de Centres Hospitaliers Impliqués Dans la surveillance Epidémiologique et la réponse aux Emergences).

Il rassemble dès à présent 26 CHU. La durée du projet est de 4 ans afin de créer un nouveau système de surveillance épidémiologique hospitalier pérenne basé sur l'utilisation secondaire des données de santé.

DESCRIPTION DU POSTE

Missions

Pour ne pas alourdir le texte, nous nous conformons à la règle qui permet d'utiliser le masculin avec la valeur de neutre.

La personne recrutée est rattachée à la Direction de la DATA et exerce ses activités de manière transversale en lien avec les autres directions de l'agence (DMI, DMNTT, DAF, DSI) et les partenaires du consortium du projet (CHU, Bordeaux Population Health Research Center, Health Data Hub, EHESP).

La personne recrutée participera activement à l'évolution et au suivi de la centralisation des indicateurs du projet Orchidée, notamment à travers la conception, le développement et l'amélioration continue d'un Dashboard dédié. Elle assurera la collecte, la consolidation et l'analyse des données en étroite collaboration avec les data scientists de la DATA et la DSI, tout en garantissant la qualité et la cohérence des informations. En lien direct avec les partenaires du projet — les CHU associés, le Health Data Hub et le Bordeaux Population Health Research Center — elle contribuera à l'identification et à la mise en œuvre des évolutions du système, qu'elles soient fonctionnelles, techniques ou méthodologiques. Ses responsabilités incluront l'animation des échanges avec les parties prenantes, la rédaction de spécifications et de documentations, ainsi que la proposition d'améliorations pour optimiser l'outil et son utilisation.

Activités

Sous l'autorité du directeur de la DATA, la personne recrutée sera en charge des activités suivantes :

- Construire et maintenir une base de données structurée, consolidant les indicateurs issus des CHU partenaires.
- Automatiser les processus de mise à jour du jeu de données centralisé pour garantir la fraîcheur et la disponibilité des données.
- Garantir la qualité, la cohérence et la traçabilité des données via des processus de nettoyage et de validation.
- Réaliser des analyses centralisées des indicateurs clés du projet Orchidée.
- Concevoir, déployer et maintenir des outils de reporting automatisés (rapports, dashboards) pour restituer les résultats et leur mise à disposition en open data.
- Participer activement aux réunions de travail et aux ateliers techniques avec les parties prenantes (CHU, Health Data Hub, Bordeaux Population Health Research Center) pour identifier les besoins, recueillir les retours et prioriser les évolutions du système.
- Rédiger des notes méthodologiques détaillant les traitements réalisés, les choix techniques et les bonnes pratiques pour assurer la reproductibilité, la transparence et la continuité de l'activité.
- Participer à la valorisation scientifique des travaux du projet (rédaction d'articles, communications, participation à des conférences) et contribuer à la diffusion des résultats auprès de la communauté scientifique.

Ces activités s'intègrent dans un environnement technique dynamique et collaboratif, mobilisant des outils de développement modernes, des langages adaptés à la science des données, et des infrastructures de calcul performantes. Le data scientist évoluera au sein d'une équipe pluridisciplinaire, en interaction étroite avec des

épidémiologistes, data scientists, statisticiens, ingénieurs et membres de la DSI ainsi que le RSSI.

Les principaux outils et technologies mobilisés incluent :

- Langages : R, Python ;
- Visualisation : Framework Shiny (R et Python), Superset ;
- Environnement collaboratif : GitLab (versionning, intégration continue, gestion des issues) ;
- Automatisation et orchestration : Apache Airflow (déploiement, supervision des workflows) ;
- Environnements de développement : VS Code, RStudio, IA Mistral.

PROFIL RECHERCHE

Niveau	Bac +5/ Bac +8 dans le domaine du traitement des données, de l'informatique ou des statistiques/data sciences.
Diplôme	
Expérience	Une expérience professionnelle d'au moins deux ans en tant que Data Scientist est requise. Une expérience dans le secteur de la santé publique, de l'épidémiologie ou des données hospitalières sera considérée comme un atout.
Aptitudes et Compétences	<p>Il est attendu pour ce poste :</p> <ul style="list-style-type: none">• Très bonne maîtrise de la programmation sous Python et/ou R ;• Solide maîtrise des pipelines de données, de l'automatisation des traitements et des méthodes de traitement de données ;• Bonnes connaissances des outils de reporting et de DataViz modernes (Shiny, Quarto, Superset ...) ;• Bonnes connaissances des outils d'automatisation et orchestration (Apache Airflow, Gitlab ...) ;• Bonne maîtrise de l'anglais (lu, écrit, parlé) ;• Rigueur scientifique ;• Très bonnes capacités d'analyse ;• Curiosité intellectuelle, esprit d'initiative et autonomie ;• Qualités relationnelles, aptitudes élevées à la discussion et au travail d'équipe ;• Réactivité, adaptation et organisation.

POUR POSTULER

Adresser les candidatures (lettre de motivation + CV) en indiquant la référence de l'annonce par courriel :

recrut@santepubliquefrance.fr

