

## PROFIL DE POSTE

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

Poste ouvert aux personnes en situation de handicap

### POSTE

|                           |  |
|---------------------------|--|
| <b>Poste</b>              | Responsable du pilotage de l'activité Data Science & Data Ingénierie (f/h)   |
| <b>Catégorie d'emploi</b> | 1  |
| <b>Emploi repère</b>      | Chargé de projet et d'expertise scientifique en santé publique (f/h)   |
| <b>Type de contrat</b>    | CDI de droit public à plein temps  |
| <b>Rémunération</b>       | 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.  |
| <b>Motif</b>              | Création suite à un redéploiement de poste.  |
| <b>Prise de fonction</b>  | Dès que possible.  |
| <b>Renseignements</b>     | Sur le poste : Christophe Bonaldi, Responsable Unité AMETIS : <a href="mailto:Christophe.bonaldi@santepubliquefrance.fr">Christophe.bonaldi@santepubliquefrance.fr</a> ,<br>Matthieu Hanf, Directeur DATA : <a href="mailto:matthieu.hanf@santepubliquefrance.fr">matthieu.hanf@santepubliquefrance.fr</a> .<br>Sur le recrutement : Marie-Christine Morel, Chargée du recrutement à la DRH – tél : 01 41 79 68 28 |

### LOCALISATION GEOGRAPHIQUE

|                |  |
|----------------|--|
| <b>Adresse</b> | Le siège de l'agence est situé à Saint-Maurice (94).<br>L'agence dispose d'implantations régionales auprès des agences régionales de la santé.<br>Localisation du poste : siège de l'agence<br>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. Elle 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)

La DATA assure un appui transversal à l'ensemble de l'agence pour 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 » (ABISS), unité « Appui et méthodes pour les études et investigations dans le domaine de la surveillance » (AMETIS), et unité « Enquêtes »). Elle intervient sur l'ensemble du cycle de vie des données de santé.

Ses missions couvrent notamment 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.

### Unité

L'unité AMETIS (Appui et Méthodes pour les Études et Investigations dans le domaine de la Surveillance) met en œuvre des méthodes quantitatives avancées et des outils d'ingénierie / science des données pour accompagner les études et dispositifs de surveillance de Santé publique France. Elle couvre l'intégralité du cycle des données, avec une expertise structurée autour de cinq axes principaux :

- Conception d'études : élaboration des designs méthodologiques et des protocoles de recueil de données.
- Ingénierie des données : nettoyage, structuration et automatisation de pipelines, ainsi que gestion de bases de données complexes.
- Modélisation et science des données : analyses statistiques avancées, statistique spatiale, métrologie et science des données.
- Visualisation et outils interactifs : Développement de cartographies dynamiques, de tableaux de bord interactifs et de rapports automatisés.
- Veille et innovation : Intégration de méthodes émergentes, IA et collaborations académiques.

En parallèle, AMETIS contribue aux dispositifs nationaux (ex : portail open-DATA Odissé) en assurant la qualité, la traçabilité et la reproductibilité des données et analyses produites.

## DESCRIPTION DU POSTE

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

### Contexte

La révolution du numérique en santé, accélérée par la crise de la Covid-19 et les préparatifs des Jeux Olympiques, a profondément modifié les fonctionnements de l'agence et plus particulièrement de la DATA. Dans un environnement où l'instantanéité des informations et les avancées de l'intelligence artificielle transforment les attentes, la DATA se doit désormais de maîtriser l'ensemble du cycle des données, de leur collecte à leur diffusion, tout en assurant une sécurité optimale et le respect du RGPD.

Acteur clé de la production d'indicateurs, la DATA met ainsi en œuvre des méthodes innovantes pour collecter, traiter, analyser et visualiser des volumes toujours plus importants de données. Grâce à son expertise, elle génère au quotidien des indicateurs fiables et précis, disponibles à toutes les échelles géographiques. Cette démarche s'appuie sur l'automatisation des traitements, la modélisation avancée et la restitution sous forme de rapports ou de tableaux de bord, afin de répondre aux exigences accrues de réactivité et de précision en surveillance sanitaire.

### Missions

Sous la responsabilité du responsable d'unité, le responsable du pilotage de l'activité Data Science & Data Ingénierie travaille en étroite collaboration avec la direction et les responsables d'unité de la DATA. Il est chargé de coordonner les travaux en ingénierie et science des données au sein de la direction.

### Activités

**Les missions du poste s'articulent autour des trois axes suivants :**

1. **Piloter, coordonner et faire évoluer les travaux en ingénierie et science des données (50 %) :**
  - Piloter les travaux en cohérence avec les orientations stratégiques de l'agence.
  - Animer les échanges transverses entre les trois unités de la direction sur les thématiques « production » via l'organisation et la facilitation de réunions internes thématiques (DATA-PROD,) et en lien étroit avec les groupes thématiques DATA-LAB, DATA-STAT et DATA-SNDS.
  - Optimiser l'organisation du travail et la répartition des tâches selon les besoins opérationnels et les compétences disponibles en lien avec les responsables d'unité des agents (AMETIS, ABISS, ENQUETE).
  - Planifier et suivre la programmation annuelle des activités, en assurant leur alignement avec les priorités institutionnelles.

- Assurer l'interface avec les directions métiers pour la définition des besoins et des calendriers, ainsi qu'avec la DSI pour la coordination technique des projets.
- Coordonner la maintenance et l'évolution des outils couvrant l'ensemble du cycle de la donnée : préparation, exploration, stockage, analyse, restitution.
- Superviser l'automatisation des processus de production et de diffusion des données (pipelines, API, exports).
- Organiser l'utilisation de GitLab comme plateforme centralisée pour le suivi, la gestion des versions et la structuration des programmes et scripts développés au sein de la DATA, dans une démarche intégrant les principes de sécurité, de traçabilité et de transition vers une approche DevOps.
- Élaborer, formaliser et maintenir une documentation structurée des procédures internes, garantissant la traçabilité des traitements et le contrôle de la qualité des données.
- Contribuer à la veille scientifique et à la diffusion des connaissances en ingénierie et science des données.

## 2. Développer des outils et méthodes d'ingénierie et science des données (40%)

- Identifier, acquérir et intégrer de nouvelles sources de données pertinentes, en réponse aux besoins de surveillance ou d'analyse.
- Assurer le nettoyage, la structuration et la préparation des données, en garantissant leur qualité, leur fiabilité, leur traçabilité et leur conformité aux exigences réglementaires.
- Concevoir, développer et optimiser des méthodes statistiques et des approches en science des données, allant de l'analyse descriptive à la modélisation avancée (ex. : détection de signaux, traitement automatique du langage).
- Définir, implémenter et interpréter des indicateurs de surveillance adaptés aux enjeux métiers.
- Concevoir, développer et maintenir des systèmes de gestion de données et des pipelines automatisés couvrant l'ensemble du cycle de traitement et de restitution des données.
- Développer des outils de monitoring et de visualisation des résultats : exports, API, tableaux de bord interactifs et reporting automatisé.
- Documenter les processus et garantir la traçabilité, la sécurité et l'éthique dans la gestion des données, conformément aux réglementations en vigueur.
- Participer aux réflexions stratégiques sur la maintenance et l'évolution des systèmes d'information structurants de l'agence, en collaboration avec la DSI.

## 3. Contribuer à la valorisation scientifique des travaux (10%)

- Rédiger des notes méthodologiques détaillant les traitements et les approches utilisées.
- Participer à la valorisation des résultats en collaboration étroite avec les épidémiologistes et les statisticiens.
- Représenter Santé publique France au sein d'instances nationales et internationales, dans le cadre d'échanges scientifiques ou institutionnels.
- Renforcer les partenariats scientifiques et contribuer à l'organisation d'événements de recherche (colloques, séminaires, ateliers).
- Encadrer des stagiaires de niveau Master 2, des internes et des doctorants sur les thématiques liées à l'ingénierie et à la science des données.

**Environnement de travail :** Le poste s'intègre 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. La personne recrutée 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 : Python / R / SAS
- Environnement collaboratif : GitLab (versionning, intégration continue, gestion des issues)
- Automatisation et orchestration : Apache Airflow (déploiement, supervision des workflows), Docker
- Formats et bases de données : PostgreSQL, fichiers Parquet, CSV ...
- Environnements de développement : VS Code, R studio, IA Mistral
- Traitement intensif : Apache Spark via serveurs sécurisés de calcul haute performance (en cours de construction)

**Participation aux activités transversales et en situation de crise :** La personne recrutée participera à la continuité d'activité au travers des permanences hebdomadaires et des astreintes de week-end sur le périmètre de ses activités. En outre la personne recrutée participera à certaines activités générales de l'unité et pourra participer à toute activité entrant dans les missions de l'Agence si cela s'avère nécessaire en situation d'urgence ou de crise.

## PROFIL RECHERCHE

|                                     |  |
|-------------------------------------|--|
| <b>Niveau<br/>Diplôme</b>           | Bac +5 à Bac +8 dans le domaine du traitement des données, de l'informatique ou des statistiques/data sciences.<br>Formation scientifique en santé publique, épidémiologie souhaitée.  |
| <b>Expérience</b>                   | Une expérience professionnelle d'au moins 5 ans en tant que Data Engineer/Scientist ainsi qu'une très bonne connaissance du secteur de la santé publique/épidémiologie est indispensable. Expérience exigée dans la gestion de projet data.  |
| <b>Aptitudes et<br/>compétences</b> | <ul style="list-style-type: none"><li>• Très bonne maîtrise de la programmation sous Python et/ou R.</li><li>• Solide maîtrise des pipelines de données, de l'automatisation des traitements et des méthodes de modélisation de données avancés.</li><li>• Solides connaissances en bases de données relationnelles et en programmation SQL (PostgreSQL).</li><li>• Solides connaissances des outils de reporting et de DataViz modernes (Shiny, Quarto, Superset ...).</li><li>• Solides connaissances des outils d'automatisation et orchestration (Apache Airflow, Gitlab ...).</li><li>• Bonnes Connaissances de l'environnement LINUX.</li></ul> <p>Il est attendu pour ce poste les compétences suivantes :</p> <ul style="list-style-type: none"><li>• Conduite de projet ;</li><li>• Management transversal ;</li><li>• Prise de décision, leadership et force de proposition ;</li><li>• Capacités d'animation, de pédagogie et de négociation ;</li><li>• Qualités relationnelles, aptitudes élevées à la discussion et au travail d'équipe ;</li><li>• Réactivité, résistance à la pression ;</li><li>• Curiosité intellectuelle, esprit d'initiative et autonomie ;</li><li>• Expression écrite et orale ;</li><li>• Rigueur scientifique.</li></ul> |

## POUR POSTULER

Adresser les candidatures (lettre de motivation + CV) en indiquant la référence de l'annonce par courriel : [recrut@santepubliquefrance.fr](mailto:recrut@santepubliquefrance.fr)