

VAGA BOLSISTA PÓS - DOUTORADO

O PROADI VERACIS convida pesquisadoras(es) para participarem do processo seletivo para a posição de Bolsista Pós-Doutorado.

Esta vaga é afirmativa e visa promover a inclusão de grupos historicamente sub-representados, buscando garantir a diversidade e a equidade de oportunidades.

Inscrições: Envie seu currículo para renata.proa@einstein.br ou via DM até 15/05/2025.

Contextualização:

Compreender o impacto de fatores ambientais e climáticos na saúde da população, em particular da população negra, é fundamental. As mudanças climáticas representam desafios significativos para o alcance dos Objetivos de Desenvolvimento Sustentável da ONU (*Sustainable Development Goals* - SDGs) até 2030 no que diz respeito à saúde materna e neonatal. Os ODS são uma coleção de 17 metas globais estabelecidas pela ONU em 2015, abrangendo temas como pobreza, fome, saúde, educação, igualdade de gênero e mudanças climáticas. O presente projeto propõe o desenvolvimento de um modelo de dados integrado, para viabilizar estudos em ciência de dados nos domínios de Clima e Saúde. O modelo buscará superar desafios de integração de dados e possibilitar pesquisas em larga escala. A proposta será avaliada utilizando um conjunto de dados composto, criado a partir da integração de dados já existentes de ambos os domínios. O plano de pesquisa inclui o estudo de modelos comuns já existentes, como o OMOP e o FHIR, a aplicação do modelo OMOP a dados do SINASC e SIM, a realização de uma análise de privacidade e a construção do modelo comum e em conjuntos de dados climáticos selecionados. A validação será realizada por especialistas, com aplicação em amostras de dados reais e avaliação do desempenho do modelo em cenários de big data. O modelo deverá estar alinhado com os SDGs 3 (Saúde e Bem-Estar) e 13 (Ação Contra a Mudança Global do Clima) e estará, em particular, focado em permitir a análise do impacto de variáveis de clima, saúde e poluição na população negra. Ao fornecer um modelo integrado de dados para Clima e Saúde, os resultados deste projeto poderão apoiar pesquisadores e formuladores de políticas públicas no desenvolvimento de estratégias de mitigação de impactos.

Objetivo da posição:

Contribuir com o desenvolvimento e manutenção de uma infraestrutura de dados robusta, segura e escalável para integrar múltiplas fontes de dados (sensores ambientais, imagens de satélite, registros de saúde e dados coletados em campo), com o objetivo de monitorar em tempo real os impactos ambientais sobre a saúde da população brasileira, com foco especial na população negra.

Atividades previstas:

- Projetar e implementar pipelines de dados espaciais e temporais (ETL/ELT).
- Trabalhar com dados do SUS, sensores ambientais e imagens de satélite.
- Integrar diferentes bancos de dados relacionais e não-relacionais.
- Apoiar a construção de painéis de visualização interativos com dados atualizados em tempo real.
- Garantir segurança, governança e rastreabilidade dos dados.
- Contribuir para documentação técnica e reprodutibilidade do sistema.
- Colaborar com cientistas de dados, profissionais de saúde pública e pesquisadores em clima.

Requisitos:

- Doutorado completo em Computação, Engenharia da Computação, Ciência de Dados, ou áreas afins.
- Experiência comprovada em engenharia de dados, incluindo construção de pipelines e integração de dados complexos.
- Proficiência em Python e/ou outras linguagens voltadas a dados (ex: SQL, Spark, etc.).
- Experiência com cloud computing (preferencialmente AWS).
- Conhecimento em bancos de dados relacionais (ex: PostgreSQL/PostGIS) e não-relacionais (ex: MongoDB, BigQuery).
- Familiaridade com APIs e integração de sensores/IoT.
- Interesse ou experiência anterior em projetos de saúde pública, clima ou justiça ambiental.
- Comprometimento com ciência aberta, reprodutibilidade e ética em dados.

Experiências Desejáveis:

- Experiência com dados espaciais e temporais, utilizando bibliotecas como GeoPandas, Rasterio, xarray ou Google Earth Engine.
- Vivência com processamento de imagens de satélite (ex: Sentinel, Landsat, MODIS) e uso de metadados georreferenciados.
- Experiência prévia com bases de dados do DataSUS (ex: SIH, SIM, SINAN, CNES), incluindo limpeza, padronização e linkage.
- Conhecimento em arquiteturas orientadas a microsserviços e pipelines modulares.
- Experiência com ferramentas de orquestração de workflows como Apache Airflow, Luigi ou Prefect.
- Familiaridade com infraestruturas de dados em tempo real, como Apache Kafka ou Spark Streaming.
- Experiência com containerização e orquestração (Docker, Kubernetes).
- Conhecimento em DevOps, CI/CD e automação de infraestrutura de dados.
- Conhecimento em visualização de dados (ex: Dash, PowerBI, Tableau).
- Familiaridade com princípios de equidade e uso de dados para políticas públicas.
-

Dedicação/Carga horária: TBD

Duração: 1 ano - início imediato (junho/25)

Junte-se a nós e contribua para o fortalecimento das políticas públicas de saúde no Brasil, promovendo uma pesquisa mais inclusiva e representativa!

[hashtag#OPORTUNIDADE](#)