



DATA SCIENTIST

The African Population and Health Research Center (APHRC) is a leading Africa-based, African-led, international research institution headquartered in Nairobi, Kenya. APHRC conducts policy-relevant research on population, health, education, urbanization and related development issues in sub-Saharan Africa.

APHRC seeks to recruit a **Data Scientist** to work on a large-scale Catalyst initiative whose goal is to transform the Research and Development (R&D) ecosystem in Africa.

Catalyst Project

This initiative aims to increase in the number of African universities and research institutes that capably test, and adopt evidence and innovations generated from health, agriculture, economics research as global public goods to achieve health impact. This will ultimately result in a strengthened ecosystem for Implementation Research in Africa, anchored on a network of highly capacitated academic and research institutions on the continent.

Duties/Responsibilities

The Data Scientist will work with consultants and APHRC's data science group in designing and build a research matching platform that links universities/research institutes based in all the regions of Africa and funding agencies for research. This will include:

- Building WebCrawler's and Scrubbers;
- Building databases;
- Indexing of databases utilizing Machine Learning/Artificial Intelligence models;
- Design and develop machine learning algorithms;
- Discover, design, and develop analytical methods to support novel approaches of data and information processing;
- Design and develop novel algorithms;
- Perform explanatory data analyses;
- Generate and test working hypotheses;
- Prepare and analyze data and identify patterns;
- Using graphics to visualize data using various visualizing tools;
- Building network analysis models;
- Building a front-end website for the research matching platform;
- Build and integrate a multilingual AI-powered chatbots using natural language processing (NLP);

Other responsibilities:

- Conduct tracking and sourcing of sector specific R&D data from across Africa for updating a regional data repository;
- Develop an interactive data platform complete with a data archiving protocol/architecture, programming the protocols, data cleaning etc.;
- Create automated systems for data harmonization and consistency checks;
- Data cleaning and harmonization including conducting all the required consistency checks for data from various partners and studies;
- Contribute to the development of scientific data standards;

- Documentation of data sets from various partners according to agreed guidelines on data documentation, archiving and sharing; and
- Facilitate discussions on data extraction, access and sharing protocols with project partners.

Qualifications, Skills, and Experience

- PhD in Computer Science, Data Science with experience in Artificial Intelligence, Machine Learning, or related technical field. Holders of Master's and Bachelor's degrees with relevant experience will also be considered.
- Experience in Software Engineering.
- Experience with one or more of the following: Natural Language Processing, text understanding, classification, pattern recognition, recommendation systems, targeting systems, ranking systems or similar.
- Experience with relevant technologies (e.g., TensorFlow, Flume, machine learning libraries).
- Relevant professional experience with applied data analytics and predictive modeling.
- Working experience in Software development, AL or ML for at least 3 years.
- In depth knowledge of Apache Tomcat/Open Source
- Excellent project and time management skills.
- Outstanding problem solving, analytical skills and attention to detail.
- Experience in data management using relational databases such as MySQL or other is essential.
- Experience in data documentation using software such as NESTAR or other.
- Experience in data analysis using STATA, R, SPSS or other statistical software.
- Knowledge of a programming language such as Python, Perl, R, JAVA, or equivalent.
- Experience working with longitudinal studies and handling large data sets is desirable.
- Experience in visualizing/presenting data for stakeholders and the use of relevant data science tools kit.
- Ability to regularly report and synthesize emerging information.
- Good written communication and presentation skills.
- Bilingual language skills (French and English) will be an added advantage.

Interested candidates are encouraged to apply through our recruitment portal <https://aphrc.org/vacancies/> **Applications will be considered on a rolling basis until the position is filled.** Only shortlisted candidates will be contacted; shortlisted candidates will be required to have a Police Clearance Certificate. Cover letters should be addressed to:

The Human Resources Officer

African Population and Health Research Center, Inc

APHRC Campus, Manga Close, off Kirawa Road, Kitisuru

P. O. Box 10787-GPO, Nairobi

Website: www.aphrc.org

APHRC is an equal opportunity employer and is committed to the protection of vulnerable persons