

# Internship Opportunity: AI and Data Science



## Requirements in a nutshell

|            |   |
|------------|---|
| Education: | Preferably a current Master's student or recent graduate in Data Science, Computer Science, AI, or related fields with a focus on quantitative analysis and/or machine learning. Suitable candidates from other disciplines such as International Development, Economics, Econometrics, Public Health, Statistics, Mathematics, or Digital Humanities with strong quantitative skills are also encouraged to apply. Bachelor's graduates with strong quantitative skills will also be considered. |
| Languages: | Fluent in English, with excellent written and oral communication skills.  |
| Must-have: | Strong programming abilities in Python. Proficiency in data manipulation tools like Excel, Google Sheets, or STATA is highly desirable. Demonstrated interest and experience in socio-economic development work. Practical experience with AI technologies, especially in areas such as generative AI, knowledge graphs, and machine learning.  |
| Location:  | Preferably based in Amsterdam, but remote working can be discussed.   |

***This opportunity is open to all nationalities.  
East African nationals are encouraged to apply.<sup>1</sup>***

## About Laterite

Laterite is a data, research and analytics firm that helps clients understand and analyze complex development challenges. We provide high-quality research services and software solutions for social impact, focusing on five sectors: education, public health, agriculture, youth & livelihoods and urbanization & migration. Based in East Africa, we offer tested data collection systems, an expert research and analytics team and a thorough understanding of the local context. This unique combination enables us to carry out full-cycle research projects, from design to data collection and analysis.

We work with universities, think tanks, international NGOs, multilateral donor organizations, foundations, and government ministries and agencies.

Our network of offices currently includes the Netherlands, Rwanda, Ethiopia, Kenya, Uganda, Tanzania, Sierra Leone and Peru. The team brings together more than 90 full time local and international staff, as well as 1,500 enumerators across countries, in a dynamic work environment. We are proud to be a culturally diverse organization.

You can find out more about Laterite on our website at: [www.laterite.com](http://www.laterite.com)

## Job Description & Key Responsibilities

---

<sup>1</sup> Costs related to relocation or visa requirements are not covered.

This internship position is affiliated with Laterite's analytics team. You will work on the development and implementation of AI-driven models and software solutions for social impact. Projects will include building and integrating knowledge graphs using Retrieval-Augmented Generation (RAG), implementing generative data augmentation using large language model (LLM) pipelines, and data scraping for constructing models in Google Earth Engine. Additionally, you will develop AI models for automated quantitative and qualitative data analysis to enhance decision-making processes in socio-economic development.

A key part of Laterite's work involves working with large datasets, including survey and geospatial data. You will use this data to build predictive models and integrate them with AI technologies to improve the precision and effectiveness of our research outputs. You will also work on client-focused, innovative data-driven solutions in the field of development.

A non-exhaustive list of potential projects you might work on includes:

- **Intelligent knowledge hubs.** Develop knowledge hubs for social impact data. Enhance data retrieval processes by integrating automated data curation pipelines with knowledge graphs, RAG, and LLM agents. This initiative involves curating unstructured data, establishing nodes and relationships within knowledge graphs, and combining these elements with a RAG pipeline to generate comprehensive, actionable insights.
- **Automated Quantitative and Qualitative Analysis.** Develop AI models that automate both quantitative and qualitative data analysis for researchers, increasing efficiency and reducing manual data handling. These models will use machine learning techniques to interpret complex data sets and provide clear, understandable analysis.
- **Generative Data Augmentation and evaluation.** Test whether we can create Large Language Model (LLM) pipelines or fine tuned LLM models to generate synthetic quantitative or qualitative data. Evaluate the quality of this data against real data. Our approach extends to training models that can generalize to unseen data, a capability beyond traditional GANs and VAEs.
- **Google Earth Engine Models:** Implement and refine data scraping techniques to build and calibrate models within Google Earth Engine, focusing on environmental and socio-economic variables. This will help in making informed decisions based on the latest satellite and geospatial data.

This is a challenging position that will provide you with hands-on experience in various types of data (administrative, geospatial, survey, etc.), coding, research, and product development. You will work alongside our senior research and management teams, who will offer guidance and mentorship throughout. It is a very practical application of the tools and techniques you have learned in your development, econometrics, statistics, and/or data science classes.

The position will help you put your curriculum into context, give you ideas for your thesis, and hopefully also some thoughts on professional opportunities after you finish your studies.

Specifically, we would like to draw on your expertise to:

- Develop and implement AI-driven models across our key sectors: Agriculture, Education, Public Health, Youth & Labor, Urbanization & Migration.
- Scout for reliable data sources that can be leveraged to build robust classification and prediction models.

- Construct detailed documentation that supports the understanding and replication of our models and methodologies.
- Design and conduct validation tests to verify the accuracy and reliability of our models before they are deployed in real-world scenarios.

## **Profile**

Our ideal candidates are master students or recent graduates that are entrepreneurial, creative, passionate about international development and social impact, structured in their thinking and demonstrating strong analytical and quantitative skills. They will be able to effectively solve problems and adapt to changing situations.

## Requirements:

- Preferably a current Master's student or recent graduate in Data Science, Computer Science, AI, or related fields with a focus on quantitative analysis and/or machine learning
- Strong programming skills in Python
- Experience with cloud services like Anvil, and working with server environments.
- Familiarity with AI frameworks and tools.
- Excellent written and oral communication skills in English
- Good understanding and practical experience with AI technologies, particularly in the areas of data scraping, knowledge graphs, and machine learning.

In addition, we welcome:

- Proficiency in working with data manipulation tools such as Excel, Google Sheets, or STATA.
- Excellent organizational and interpersonal skills, self-motivation, and the drive to thrive in a fast-paced environment where timelines can often be unpredictable.

## **Additional information**

This opportunity is offered as a paid internship.

Dates are flexible and can be decided between Laterite and successful applicants. We expect the interns to work with us for 2 to 6 months.

## **What next?**

### **1). Verbal Reasoning and Quantitative Assessment**

The first step is to complete a 30-minute verbal reasoning and quantitative assessment for which no special preparation is needed.

Link to the assessment: <https://form.jotform.com/241503915545556>

### **2). Submit application**

Successful candidates will then be invited to upload their CV and cover letter via our online application system.

### **3). Analytical assessment**

This assessment aims to gauge your capacity to create functioning applications in Python using publicly available APIs. This exercise is similar to the type of applications you will be asked to build as part of this internship at Laterite.

#### **4). Interviews**

Successful candidates will then be invited to a first interview.

#### **Deadline**

Applications will be considered on a rolling basis. Details on rolling applications can be found on the website: <https://www.laterite.com/vacancies/>