AI Engineer - Consultant

Terms of Reference

1. Summary

The Development Impact Group's AI Lab is launching a recruitment process for an **AI Engineer** – find out more about the Lab's work [here](). They will have the opportunity to work across the different projects of the AI Lab, dependent on their experiences, interests, and project needs. The candidate hired through this process will be offered employment as soon as possible and will have the option to work remotely. To apply, please send your **CV** and **cover letter** to [dime_ai@worldbank.org]() **(with "AI Engineer – Application" in subject header**) by **31 August 2025 11:59 PM EST.**

2. Background and objectives

a. Development Impact Group (DIME)

The Development Impact Group is part of the World Bank's Development Economics Vice-Presidency (DEC). The Group's purpose is to increase the use of research evidence, in particular the use of impact evaluations (IE), in the design and implementation of public policy and to develop institutional capacity and motivation for evidence-based policymaking.

The Development Impact Group's work focuses on:
- Evaluating programs at scale to answer policy-relevant questions;
- Developing innovative data systems and data analytics tools for real-time decisions;
- Using causal inference analysis to identify mechanisms that improve results;
- Building capacity to strengthen country institutions for evidence-based policymaking.

The Group aims to overcome the challenge of identifying true cause-and-effect relations in policy programs. By linking researchers to policymakers and feeding results back into policies, we foster systematic use of evidence, which informs adoption, mid-course corrections, and scale-up of policies. With a portfolio of more than 220 impact evaluations, the Group operates across all sectors in about 60 countries spanning the globe. The IEs test a variety of interventions and mechanisms to understand why policy succeeds or fails and how to improve policy design and implementation to obtain better results.

As part of the aforementioned mission, the [AI Lab]() is pioneering the next frontier of impact evaluation, leveraging AI and machine learning to develop and implement research, interventions, and tools that address pressing global challenges. Operating with a commitment to maintaining the highest ethical standards and rigor, particularly considering vulnerable communities, the team's work has uncovered new, evidence-based approaches to predict food insecurity, mitigate hate speech and misinformation, advance conflict early warning systems, and address gender biases. Our mission is to leverage cutting-edge technology to advance social progress and sustainable development worldwide. The team seeks to employ recent technological advancements in ML, AI, and adjacent disciplines as an enabling force for the IE work the wider Development Impact Group conducts.

b. Project details

The AI Engineer (ST Consultant) can work across all three project streams.

ZeroHungerAI: Food Security threatens hundreds of millions globally, with undernourishment rising from 580 million in 2012 to 730 million in 2022 (FAO, 2024). Climate change, population growth, political instability, and economic disparities worsen the issue, hitting vulnerable communities hardest. Predicting these crises is challenging due to outdated data. DIME AI's research pilot ([research paper](), [blog article]()) uses natural language processing on non-traditional data sources, such as news articles, to detect early signs of food crises, improving forecasts by up to 46% a year in advance, allowing more time for action. This provides extra time for proactive measures to be taken. Real-time news integration aids decision-making in data-scarce environments, improving humanitarian responses and saving lives. The team is now looking into turning the successful research pilot from 37 countries into a live public good with weekly updates at the district level across 82 countries. Additionally, the project will consist of research exploration, including adding local languages to improve predictive accuracy and constructing geospatial signals to provide higher-resolution predictions. The aim of the project is also to be integrated with existing food insecurity and anticipatory action programming.

ImpactAI, [DIME AI's flagship project](), harnesses the power of large language models (LLMs) to sift through swathes of impact evaluation research and accurately extract insights that provide effortless access to causal evidence for improved decision-making. With an added generative AI layer, this tool lets users provide natural language queries and receive personalized, precise responses. Tailored for a diverse audience committed to global development and humanitarian efforts, primary users include development practitioners, policymakers, international development organizations, academic institutions, research institutes, think tanks, and the economic research community. As a one-stop source that lowers the barriers to accessing high-quality research, ImpactAI offers easy access to cutting-edge ideas, best practices, and empirical evidence, thereby enhancing the effectiveness of development projects worldwide. Learn more about the product [here]().

SocialAI: Bias, Hate Speech, and Misinformation are disproportionately and increasingly affecting marginalized groups in the Global South. Hate speech detection (HSD) models help combat harmful online content but often rely on U.S.-centric data, overlooking other English dialects and other languages worldwide. Conventional testing methods, often involving biased datasets, can also inflate content moderation performance. Culturally tailored HSD models combined with human oversight can better identify hateful content. DIME AI's previous research ([research paper]()) on Nigerian tweets shows that combining culturally tailored models with human review of just 1% of flagged tweets can moderate 60% of hateful content. Understanding social media usage helps track societal trends and improve policy responses, such as curbing interethnic violence. DIME AI uses its top-performing model to identify Nigerian users sharing hate content and targets them with celebrity-led prosocial ads. Building from this knowledge, DIME AI is now scaling this line of work to additional countries as well as developing an Entertainment Media Gender Monitor for conducting content analysis of programming consumed by adolescent girls and young women to focus on gender norms. Simultaneously, we will be evaluating context-specific gender bias exhibited by LLMs when operating in the Global South.

c. Responsibilities

The AI Engineer will coordinate their work with the project PIs and other research team members. You will collaborate closely with the domain experts and incorporate real-world constraints and validity checks into the models. Specific duties include, but are not limited to:

- **Collect and prepare large corpora for model training:** Source, clean, and organize large-scale datasets to build high-quality training corpora for domain-specific LLM development.

- **Fine-tune language models:** Support AI Scientists in adapting state-of-the-art models such as Llama, Gemini, or GPT to the specific requirements of the project through domain-specific fine-tuning on the prepared corpus.
- **Enhance and train encoder architectures for information extraction**: Design, implement, and optimize encoder-based models for the extraction of structured causal knowledge from unstructured text such as research studies or news articles.
- **Implement a robust evaluation framework:** Support AI Scientists in designing and implementing comprehensive evaluation pipelines to assess model performance using metrics such as accuracy, precision, recall, efficiency, and scalability, ensuring alignment with project objectives and data characteristics.
- **Deploy models in cloud environments:** Package and deploy the trained models in production-ready cloud infrastructure (e.g., GCP or AWS), ensuring scalability, accessibility, and low-latency performance for end-users.
- **Disseminate research findings:** Target publication in top-tier NLP and AI conferences by documenting methodologies, experimental results, and practical insights gained throughout the development cycle.

3. Requirements

- 3 years of hands-on experience in an AI/NLP/ML role focusing on methods such as encoder models, decoder models, knowledge graphs, and AI-based approaches to forecasting
- Master's degree in Machine Learning, Computer Science, Statistics, or a related field
- Strong proficiency in Python, with experience in ML frameworks (PyTorch, TensorFlow, scikit-learn)
- Experience with large language models (incl. training them) and traditional machine learning methods
- Experience with cloud platforms (AWS, GCP, Azure) and containerization (Docker, Kubernetes)
- Knowledge of MLOps practices, model deployment, and monitoring
- Familiarity with big data ecosystems and parallelized computing methods
- Motivated, self-starter, able to work with little supervision, with demonstrated ability to work effectively and sensitively in teams

4. Preferred Qualifications

- 4+ years of hands-on experience in an AI/NLP/ML role focusing on methods such as encoder models, decoder models, knowledge graphs, and AI-based approaches to forecasting
- PhD degree in Machine Learning, Computer Science, Statistics, or a related field
- Excellent analytical and communication skills to convey findings to non-technical audiences
- Experience working with humanitarian organizations or development agencies
- Publications or conference presentations on AI and machine learning

5. Contract arrangements

The AI Engineer will be hired as a short-term consultant (STC) starting as soon as possible. They will be able to work remotely or in person in Washington, DC. Onboarding will include access to internal World Bank systems and technical support as needed.

6. Intellectual Property

The World Bank shall, solely and exclusively, own all rights in and to any work created in connection with this agreement, including all data, documents, information, copyrights, patents, trademarks, trade secrets or other proprietary rights in and to the work. The AI Engineer is not allowed to post or publish (electronically or in print) any project-related information without the explicit permission of the Impact Evaluation Team.

7. Apply now

To apply, please send your **CV** and **cover letter** to dime_ai@worldbank.org by **31 August 2025, 11:59 PM EST.** Only shortlisted candidates will be contacted.

We are proud to be an equal opportunity and inclusive employer with a dedicated and committed workforce, and do not discriminate based on gender, gender identity, religion, race, ethnicity, sexual orientation, or disability. Learn more about working at the World Bank and IFC, including our values and inspiring stories.