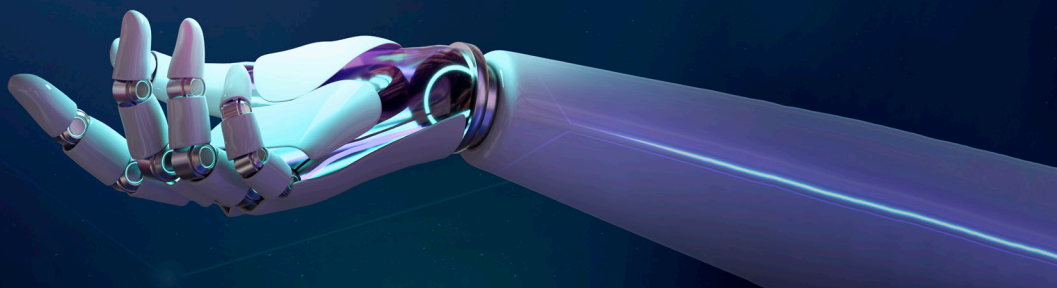AFRICAN SCHOOL OF EVALUATION
EDITION 1

AI for Transformative Evaluation: Practical Guide on Applications of AI in International Development

Provided by 3ie

PROPOSED TRAINING

Title:	AI for Transformative Evaluation: Practical Guide on Applications of AI in International Development
Facilitators:	Fiona Kastel, Sanchi Lokhande (In-person)
Workshop Length:	5 days
Audience Level:	All levels, with limited to intermediate experience using AI
Approximate Group Size:	30-35

This beginner to intermediate-level training will empower evaluators to leverage AI tools effectively in international development. Attendees will gain foundational knowledge in AI concepts, guidance on using AI across development contexts, and hands-on experience with large language models (LLMs) to support decision-making and analysis in program evaluation

SUGGESTED TRAINING OUTLINE

Day 1: Introduction to AI for Evaluators

Theme : Understanding foundational concepts and building curiosity about AI's role in development evaluation

Goals: Introduce key terms and ideas, demystify how AI works, and encourage critical engagement with AI tools

1. Introductions and Setting the Stage: Thinking Critically About AI
- Icebreaker activity and discussion: Guess if content is human- or AI-generated
- Highlights of recent AI developments and applications in the news
- Discussion: The role of AI in our lives and work

2. Foundations of AI: Introduction to Core Concepts in AI
- Overview of Artificial Intelligence, Machine Learning, Generative AI, and Large Language Models
- Evolution of AI and the current landscape
- Peer-led reflection: Where have you already encountered these technologies in your work?

3. Focus on Generative AI tools
- Introduction to key GenAI chatbots (e.g., ChatGPT, Claude, Gemini)
- Hands-on activity: Compare responses using LLM Arena
- Discussion on differences in outputs and potential use cases

4. Hands-On Working Session: Critical Evidence Exercise
- Group activity comparing AI-generated and traditional research workflows
- Short presentations and facilitated reflection

5. Optional Sandbox Time: Participants can explore a tool introduced, try a use case from their work, or request help from facilitators

Day 2 : Introduction to AI in Evaluation Workflows

Theme: Exploring how AI can support evaluation thinking

Goals: Introduce key entry points for AI across the evaluation cycle, explore use cases, and build a practical AI tool

1. Recap

2. AI across the Evaluation Cycle
- Overview of how AI intersects with different phases of the program cycle
- Discussion: Where can AI fit into evaluation work?
- Use-cases from development settings

3. Hands-On Working Session: Build a Custom Chatbot
- Guided exercise using a no-code platform to build a task-specific chatbot
- Participant-led presentation and discussion of custom chatbots

4. Optional Sandbox Time: Continue refining chatbot, apply it to an evaluation workflow, or consult facilitators on implementation ideas

Day 3: Applying AI in Evaluation

Theme: Exploring how AI can support evaluators in the early stages of designing and planning an impact evaluation.

Goals: Use LLMs to support early evaluation thinking, including ideation around theories of change, research questions, and evaluation design.

1. **Recap**

2. **Quick Overview: What is Impact Evaluation?**
   - Core concepts: causality, counterfactuals, theory of change, and designs
   - Discussion: What challenges do you face in designing evaluations? Could AI support this?

3. **Hands-On Working Session: LLMs for Early Evaluation Thinking**
   - Guided exercise using LLMs to brainstorm theories of change, research questions, and research designs

4. **Scenario-Based Task: Respond to a ToR and design an initial approach using an LLM**

5. **Optional Sandbox Time: Refine your theory of change, try a use case from their work, or request help from facilitators**

Day 4: Applying AI in Evaluation (Pt.2)

Theme: Exploring advanced applications of AI in data collection, measurement, and analysis

Goals: Introduce participants to frontier AI techniques that enhance evaluation practices and provide hands-on experience with LLMs for coding and analysis

1. **Recap**

2. **Advanced AI Applications in Evaluation**
   - Using web crawling and scraping for context mapping and secondary data collection
   - Machine learning and computer vision for measuring outcomes and covariates using satellite or image data
   - Introduction to causal trees and causal forests for assessing treatment effect heterogeneity
   - Brief case examples from evaluation practice
   - Group Reflection: Which frontier application is most realistic for you? Why? What other applications have you heard of?

3. **Hands-On Working Session: LLMs for coding and analysis of data**
   - Guided exercise using LLMs to generate and refine code in Stata, R, and JavaScript
   - Discussion on evaluating outputs, debugging with AI, and maintaining analytical transparency

4. **Optional Sandbox Time: Try coding a basic data cleaning or visualization task from your own dataset**

Day 5: AI and Responsible Use

Theme: Exploring AI for document-heavy workflows and reflecting on responsible and ethical integration

Goals: Use LLMs to support text-based tasks; critically examine the risks and ethical dimensions of using AI in evaluation; and synthesize key learnings from the training.

1. **Hands-On Working Session: AI for Text-Based Tasks**
   - Guided exercise using LLMs for summarization, literature review support, and qualitative analysis
   - Examples of how AI can help structure evidence synthesis and accelerate document-heavy workflows

2. **Ethics, Risks, and Responsible Use of AI**
   - Discussion on AI limitations, hallucinations, and risks in evaluation contexts
   - Issues of bias, transparency, accountability, and data protection
   - Best practices and safeguards for responsible AI use
   - Emerging guidance and standards from the development and evaluation communities
   - Scenario-Based Decision-Making: Prompts exploring ethics in action

3. **Key Takeaways and Reflections**
   - Review of key concepts and tools discussed
   - Participant reflections on what surprised, challenged, or excited them
   - Final Peer Roundtable: How will you apply one new insight in your work

4. **Optional Sandbox Time: Revisit any tool or concept from the week with facilitator guidance**