



## Workshop Syllabus

### AI + Gender Data Lab Capacity Building Program in Rwanda

#### INSTRUCTOR

Dr. **Shu FU** ([fushu@sjtu.edu.cn](mailto:fushu@sjtu.edu.cn)) is an assistant professor in the School of International and Public Affairs at Shanghai Jiao Tong University (SJTU), China. He received his PhD degree in Political Science at the University of Chicago. He specializes in American politics and empirical methodology. His research is published at top journals in political science, including *Journal of Politics*, *World Politics*, *Political Research Quarterly*, and *Presidential Studies Quarterly*. He teaches courses on Introduction to American Politics, Causal Inference, and Social Science Inquiry, and was awarded the 1st Prize in the Young Faculty Teaching Competition at SJTU.

#### FACILITATORS

Zhoutong WAN  
Shiyu HAN  
Mingwei LI

Yuchen JIANG  
Xinyue XIE  
Boyue WANG

To ensure the skills transferred in the workshop is consolidated, the participants will be paired with a facilitator during the workshop, to follow up on the use of AI in making metadata more accessible. Each facilitator is expected to collaborate with the NISR participants in solving analytical or coding problems, ensuring each participant can produce the targeted report or presentation.

#### ORGANIZER

Mercedes FOGARASSY is a policy analyst at PARIS21 (Partnership in Statistics for Development in the 21st Century), hosted at the OECD. She works at the intersection of data, innovation and capacity development, supporting national statistical systems in low- and middle-income countries in becoming AI-ready. Before joining PARIS21, she advised governments and international organisations on data ethics and citizen engagement at the global data collection company and on information integrity and digital policy at the OECD. Originally from Vancouver, Canada, Mercedes holds degrees from the University of Toronto, and Sciences Po Paris specialising in international relations and global risks.

## OBJECTIVE

This workshop aims to enhance the capacity of participants in using artificial intelligence (AI) tools to improve the quality and accessibility of gender-related metadata, with a focus on poverty data from Rwanda's EICV7 survey. Participants will be introduced to the general concepts and evolution of AI, including its transformative potential and associated risks in the context of data governance. The workshop will provide hands-on training in writing effective prompts for large language models (LLMs) such as ChatGPT, Deepseek, and Gemini, in order to generate well-structured and machine-readable metadata. Emphasis will be placed on using AI to translate complex statistical outputs into formats that are more user-friendly and policy-relevant. By the end of the workshop, participants will be equipped with practical skills to leverage AI for improving the discoverability and usability of gender and poverty-related data.

## SOFTWARE AND READING MATERIALS

For this workshop we will be using R. R is a programming language that is especially powerful for data exploration, visualization, and statistical analysis. You might find helpful the following resources, which are additional readings for further learning.

- *R for Data Science* by Hadley Wickham and Garrett Grolemund, which is in open access at <https://r4ds.had.co.nz/>.
- *ggplot2: Elegant Graphics for Data Analysis*, by Hadley Wickham, which is available at <https://ggplot2-book.org/index.html>.

In addition to R, this workshop introduces several AI tools designed to enhance the usability, accessibility, and machine-readability of statistical metadata:

- ChatGPT (<https://chat.openai.com>): A large language model that can help generate and refine metadata descriptions, explain data trends in plain language, and convert technical data into user-friendly narratives.
- Deepseek (<https://www.deepseek.com>): An open-source LLM optimized for structured data handling and code generation, particularly useful for automating metadata creation.
- Gemini by Google (<https://deepmind.google/technologies/gemini>): A multimodal AI tool capable of interpreting charts, tables, and visualizations, aiding in transforming statistical findings into accessible formats.
- Claude by Anthropic (<https://www.anthropic.com>): A conversational AI assistant known for its interpretability and safety features, useful for drafting, structuring, and verifying metadata and documentation with a focus on transparency.

## CONTENT ACCESS

All the workshop content can be accessed at  
<https://github.com/fushu18/RwandaWorkshop2.0>

## WORKSHOP OUTLINE

### DAY 1 – Tuesday 5 August

#### **Session 1.a(9:30–10:45) Opening & Course Overview**

Mercedes Fogarassy and Shu Fu

#### **Session 1.b (11:00–12:15) AI for Official Statistics**

Mercedes Fogarassy

#### **Session 2.a (14:00–15:15) AI Concepts and Evolution**

Shu Fu and SJTU team

#### **Session 2.b (15:30–16:45) AI Potentials and Risks**

Shu Fu and SJTU team

### DAY 2 – Wednesday 6 August

#### **Session 3.a(9:30–10:45) AI Large Language Models and Prompts**

Shu Fu and SJTU team

#### **Session 3.b (11:00–12:15) R Coding Revision: Tidy Data and ggplot2**

Shu Fu and SJTU team

#### **Session 4.a (14:00–15:15) Exploratory Session 1 – Application on EICV7**

Shu Fu and SJTU team

#### **Session 4.b (15:30–16:45) Exploratory Session 2 – Application on EICV7**

Shu Fu and SJTU team

### DAY 3 – Thursday 7 August

#### **Session 5.a(9:30–10:45) Exploratory Session 3 – Application on EICV7**

Shu Fu and SJTU team

#### **Session 5.b (11:00–12:15) Exploratory Session 4 – Application on EICV7**

Shu Fu and SJTU team

#### **Session 6.a (14:00–15:15) Communicating with AI generated Report**

Shu Fu and SJTU team

#### **Session 6.b (15:30–16:45) Communicating with AI generated Report**

Shu Fu and SJTU team