IUSSP-PAA Webinar on AI & Demography

Part 1 - 16 October 2026

Written Questions & Answers

Questions answered live (or unanswered)

2. Chido Chinogurei
Thanks so much for the presentations. My question goes to Emillio, In my experience organisations have tended to go towards not sharing data due to privacy protection fears particularly with LLMs on accessing datasets what would probably propel promoting synthetic data use to address this gap, particularly for longitudinal datasets.

3, Maxence Fréchette
Thank your for your presentations. Emilio Zagheni's fifth example mentioned AI literacy. I want to ask my question to both panelists. In your opinion, what do demographers must keep in mind when using AI in research ?

4, Andrea Lisette Aparicio Castro
As AI becomes more widely used, how do you think we could distinguish real changes in population behaviour from changes in the digital data we use to study them? What approaches might help us address this challenge?

5. Nafeesath K A
The amount of time people can use internet or have access depends on the income category, gender and the area live as well right. doesn't it have the effect on the research outcome which assumes that most have digital devices access

6, Corey Sparks
I'm very interested in synthetic data, and have used AI to make some examples of different types of administrative data with statistically representative marginal distributions. I'd love to hear from the group about their opinions on generating more useful synthetic data that preserve more types of statistical associations.

8. Hal Caswell
Hi Emilio --- just a comment. The quote from di Lampedusa reminds me of what W. Ross Ashby, one of the founders of cybernetics, called the Law of Requisite Variety. In order to keep a system constant in the face of external changes, some part of the system must have at least enough flexibility to match the variation of the external changes.

Questions answered in writing :

1. Boyam YAMEOGO
Thank you for these very insightful presentations. In some contexts, demographic data are scarce, outdated or of uneven quality. How can AI models be trained and validated reliably in such settings, and what safeguards would you recommend to avoid biased estimates?

→ Emilio Zagheni
In data-poor demographic settings, the goal should not be to make the model appear equally confident everywhere. My sense is that it should be to make its confidence proportional to the quality and representativeness of the evidence.

7. Andrea Lisette Aparicio Castro
@Mathew, I was really surprised to see that AI simulations were closer to real-world behaviour than survey responses in some of your examples. How consistently have you found this across different populations and behaviours, and what do you think explains this?

→ Mathew Hauer
Yes. We see this consistently across our simulations. There's lots of reasons why. I might say I desire to eat a healthy salad for lunch (stated preference) but I eat a greasy cheeseburger (revealed preference).