Comparative Population Studies
Vol. 50 (2025): xxvii-xxxii (Date of release: 17.12.2025)

Editorial

Reflections on Population Studies in the Age of AI*

Frans Willekens

Abstract: In his editorial, former editor of *Comparative Population Studies* (CPoS) Frans Willekens reflects on the use of artificial intelligence (AI) in population studies. Effective and responsible use of any tool requires a basic understanding of how it works, when it may be used, and when its use should be avoided. When this fundamental principle is observed, AI tools can enrich learning and research and help advance the frontiers of knowledge. Epistemic integrity and accountability remain essential; the advent of AI does not diminish that core value. Although generative AI is currently dominated by machine learning and relies on statistical inference to make predictions and generate content, rule-based AI, which dominated AI in the early days, is making a comeback. Students of population should critically engage with the expanding landscape of AI systems and resist the tendency towards technological monoculture. They should cultivate substantive collaborations with computer scientists to develop domain-specific AI systems that fully prepare population studies – with demography at its core – for the era of AI.

Keywords: Artificial Intelligence (AI) · Population Studies · Machine Learning · CPoS

1 Introduction

Congratulations to *Comparative Population Studies* (CPoS)/*Zeitschrift für Bevölkerungswissenschaft* (ZfB) and the Federal Institute for Population Research (*Bundesinstitut fur Bevölkerungsforschung*, BiB) on the journal's 50th anniversary. CPoS has positioned itself as an international and interdisciplinary journal of population studies with a strong focus on comparative research. The editorials accompanying the 50th anniversary highlight the unique editorial process, the guiding principles and the strategic considerations that have shaped the journal. The strategic decisions

* This article belongs to a series celebrating the journal's 50th anniversary.

URL: www.comparativepopulationstudies.de
DOI: https://doi.org/10.12765/CPoS-2025-17
URN: urn:nbn:de:bib-cpos-2025-17en

included the publication online, open access at no cost to authors and readers, publication of accepted articles as soon as they are ready (rolling publication model), and a publication support unparallelled by most other population studies journals. CPoS has the comparative advantage of being a marvelous ecosystem of authors, reviewers, publisher and editorial staff (including language and technical editors) supporting the publication process from start to finish. Previous editorials documented the impressive trajectory of CPoS from its foundation in 1975 to the internationally recognised academic journal it is today. In a previous editorial, I had the opportunity to share my perspective on CPoS (*Willekens et al.* 2023).

The journal's transition to online publication reflects a recognition of the opportunities offered by digital innovation. Today, artificial intelligence (AI) presents another transformative frontier. AI is the outcome of decades of curiosity-driven and mission-driven research in which scientists systematically analyzed biological, cognitive and social processes and used simulation to reproduce the processes *in silico* to enhance our understanding of how humans reason, learn and make decisions, how populations evolve and societies function and why they can derail. Contemporary knowledge has advanced to a stage at which several cognitive functions can be simulated accurately or approximated with a high degree of reliability when supplied with sufficient high-quality data and appropriate statistical models and techniques. That knowledge is now used to *build* systems that exhibit cognitive capabilities that are traditionally associated with humans, such as reasoning and learning. In this editorial, I reflect on population studies in this new era. First, a few notes on AI.

2 AI

AI is an interdisciplinary field of research in modelling and simulation aimed at building computational systems that perform tasks efficiently and effectively with minimal human intervention and in a variety of situations (*Russell/Norvig* 2022). AI systems are capable of reasoning and learning. Two forms of reasoning are traditionally distinguished. The first is deductive reasoning, which applies formally defined logical rules (if … then) to propositions to derive conclusions that follow necessarily from them. Consider the proposition "two individuals are siblings". If the two individuals have at least one biological parent in common and the two individuals are not the same person (the logical rules), then the proposition is true. The second is inductive reasoning, which generalises from observations. Generalisation may point to logical rules or causal dependencies, but can never be certain. The two forms of reasoning underly the two main paradigms in AI. *Symbolic or rule-based AI* systems use primarily deductive reasoning, applying formally defined rules to known facts (data) to derive conclusions that logically follow (logical inference). It was the dominant paradigm in the early stages of AI from the mid-1950s to the mid-1980s. *Machine-learning (ML)*, which currently dominates AI, relies on inductive generalisation. ML uses statistical inference to identify and model patterns in data. Inferring patterns from unstructured data without any

prior information is unsupervised learning, while inferring patterns in structured (e.g. tabular) data is supervised learning. Machines or computational systems can also learn from examples (imitation learning) and the positive or negative feedback it receives (reinforcement learning). Much of current research involves combinations of different learning paradigms (*Jordan/Mitchell* 2015: 258). Since ML-based systems learn from statistical associations, they are sometimes referred to as *statistical AI*. Learning from statistical associations is the dominant mode of learning in AI systems today. Interestingly, learning from statistical associations is also the dominant mode of knowledge acquisition in population studies as we know it today, thanks to the user-friendly statistical software packages.

The methods of statistical inference used in ML range from linear regression to Hidden Markov Models to Deep Neural Networks (DNNs). A DNN is a neural network consisting of nodes (neurons) organised in multiple layers and connected by weighted edges (*Bishop/Bishop* 2023; *Russell/Norvig* 2022: Chapter 22). Deep neural networks learn multiple levels of increasingly abstract patterns from very large datasets, a process known as Deep Learning. The presence of multiple hidden layers confounds the information flow through the network. It explains why DNNs are often designated as *black boxes* and criticised for lack of transparancy. For a more extensive introduction to AI systems and their use in academic research, see *Willekens* (forthcoming).

3 Population studies in the age of AI

AI systems incorporate traditional software tools such as word processors, statistical packages, and search engines. Effective and responsible use of any tool requires a basic understanding of how it works, when it may be used, and when its use should be avoided. AI systems, including Large Language Models (LLMs), have built-in limitations. Common limitations include hallucination, "out of distribution" (OOD) prediction, prediction based on incomplete and defective data retrieved mainly through web-scraping, and the opaqueness of the DNN on which LLMs rely. These limitations have been documented and procedures exist for reducing their negative effects. Except for hallucination, these limitations are familiar to most of us. Population scientists have been trained to deal with data deficiencies, OOD extrapolation and lack of transparency of the documentation of models and data. That expertise can be used to evaluate AI systems. A comparative analysis of AI systems reveals significant differences. For instance, my experience tells me that ChatGPT[1] is good at generating *creative text* that prioritises imagination, emotion, entertainment, and other ways to engage the reader, while Google's Gemini[2] is better in generating *factual text* that is objective, factual, accurate, up-to-date, and

1 https://chatgpt.com/

2 https://gemini.google.com/app

verifiable. I propose a pragmatic approach to AI that seeks to capture its benefits while preventing unanticipated consequences. The approach stimulates critical thinking about AI, the role of modern technology in academic research, and the balance between innovation and preservation in our discipline.

A particular useful subject of AI-related research concerns effective querying and prompting. The user's prompt is a major factor in determining the response generated by an LLM within its built-in constraints. Since LLMs capture the meaning of words and sentences from their context (distributional semantics), it is crucial for users to augment a query with adequate context. For instance, asking ChatGPT, Gemini, Microsoft Copilot[3] or any other chatbot for the definition of fertility, it will refer to fertility as the capacity to reproduce. It is the established definition in biology and medicine. In demography, the capacity to reproduce is denoted by the term "fecundity". To get an accurate definition, the user should describe the context. In response to the following prompt "How is fertility defined in population studies?", Gemini responds: "In demographic contexts, fertility refers to the actual production of offspring, rather than the physical capability to reproduce, which is termed fecundity." It adds the source of that information, which is an advance relative to previous versions. The sources vary, however. They include wikipedia[4], grokipedia[5] and Statistics Canada[6]. In the query, the user may request the definition from a particular source. For instance, if the user wants the definition of fertility provided by the demographic dictionary Demopaedia[7] (developed by the *French Institute for Demographic Studies* (INED[8])), Gemini will return the appropriate definition if it has access to the primary source or a secondary source that includes the Demopaedia definition of fertility, e.g. lecture notes published online. The ability to design clear, concise, and effective prompts is increasingly recognised as a valuable skill. These skills should be part of any educational programme. Prompting is essentially computer programming in a natural language. The more detailed the instructions, the better the response. Users can write the prompt in any major language. The built-in Natural Language Processing (NLP) module automatically translates the request.

AI systems evolve very rapidly and consecutive releases may differ considerably. Recent versions of chatbots include retrieval augmented generation (RAG) extensions of LLMs. An LLM-RAG retrieves relevant external documents at inference time and uses the information as context to generate the text requested by the user. The prompt determines which documents are retrieved. The user may also attach documents to the prompt. Since the introduction of RAG, hallucination has

---

3 https://copilot.microsoft.com/

4 https://en.wikipedia.org/wiki/Fertility#:~:text=Demographers%20measure%20the%20fertility%20rate,over%20a%20period%20of%20decades

5 https://grokipedia.com/page/Fertility_factor_(demography)

6 https://www.statcan.gc.ca/en/concepts/definitions/previous/prefertility

7 http://en-ii.demopaedia.org/wiki/Fertility

8 https://www.ined.fr/fr/recherche/projets-recherche/S0518

decreased markedly and responses are more up-to-date. Users should recognise, however, that AI systems cannot access articles or books that are not available on the open web. The open web excludes documents and data that are published in print only or are published online but require payment or authentication. Although an AI system may not have access to the full text of an article, it may have access to the abstract and secondary sources. The selective nature of the data AI systems rely on poses a major problem in scientific research. The resulting issues of bias and representation have much in common with the challenges encountered when social media data are used in population research (*Zagheni/Weber* 2015).

Some AI systems are designed to facilitate systematic literature reviews. AI-powered search engines include Consensus[9], Elicit[10], SciSpace[11] and Perplexity[12]. The latter searches the Web in real-time. Each system is specialised in some aspects of a literature review. Search engines use external LLMs for conversation, except Google Search which uses Gemini for its AI Overview. Since the emergence of LLM-RAG and AI-powered search engines, systematic literature reviews have proliferated at an unprecedented pace in some disciplines. In response, publication venues have begun to place restrictions on such reviews (see *Castelvecchi* 2025). The usage of AI in academic research is transforming the research process, as well as the dissemination of results. Students of population should critically engage with AI. They should explore the expanding landscape of AI systems and resist the tendency toward technological monoculture.

The current AI systems used in academia are general-purpose systems. A natural evolution is toward specialisation, i.e. domain-specific AI systems. Such an evolution is not only effective, but is also a guarantee that domain experts maintain control over AI systems. In this vision, population scientists need to cultivate substantive collaborations with computer scientists to develop AI systems embedded in population studies. In the ongoing transition from monolithic AI systems to modular AI systems (Domain-Specific AI or Agentic AI), the design of a Demographic AI Agent or Population Studies AI Agent would solidify the position of population studies in the age of AI. Irrespective of whether or not a domain-specific AI system will be developed, a debate is required about population studies as a discipline in the age of AI. CPoS may stimulate that debate in ways that uphold its initial (1975) mission to be an open forum for scientific debate of population issues from an interdisciplinary perspective.

9 https://consensus.app/

10 https://elicit.com/

11 https://scispace.com/

12 https://www.perplexity.ai/

4 Final remarks

Editors of academic journals such as CPoS are guardians of quality within the scholarly community. Their role is central to ensuring that results of innovative research are disseminated and authors uphold established standards and practices, as part of our shared commitment to maintaining the independence, integrity, credibility and accountability of academic research. CPoS, being a marvelous ecosystem of authors, reviewers, publisher and editorial staff supporting authors throughout the publication process, is well positioned to foster the advancement of population studies in the age of AI. I wish everyone involved every success in the years to come.

References

Bishop, Christopher M.; Bishop, Hugh 2023: Deep learning: Foundations and concepts. Cham: Springer Nature. https://doi.org/10.1007/978-3-031-45468-4

Castelvecchi, Davide 2025: Preprint site arXiv is banning computer-science reviews: here's why. In: Nature, News article. https://doi.org/10.1038/d41586-025-03664-7

Jordan, Michael I.; Mitchell, Tom M. 2015: Machine learning: Trends, perspectives, and prospects. In: Science 349,6245: 255-260. https://doi.org/10.1126/science.aaa8415

Russell, Stuart; Norvig, Peter 2022: Artificial intelligence: A modern approach. 4th edition (Global Edition). Harlow, UK: Pearson Education Limited.

Willekens, Frans 2026: Population studies in the age of artificial intelligence. BiB.Working Paper 1/2026. Wiesbaden: Bundesinstitut für Bevölkerungsforschung.

Willekens, Frans et al. 2023: The Past, Present and Future of CPoS. In: Comparative Population Studies 48: i-vi. https://doi.org/10.12765/CPoS-2023-24

Zagheni, Emilio; Weber, Ingmar 2015: Demographic research with non-representative internet data. In: International Journal of Manpower 36,1: 13-25. https://doi.org/10.1108/IJM-12-2014-0261

Prof. Dr. ir. Frans Willekens (✉). Honorary Fellow at Netherlands Interdisciplinary Demographic Institute (NIDI). The Hague, Netherlands. Emeritus Professor of Population Studies at the University of Groningen. Netherlands. E-mail: willekens@nidi.nl
URL: https://nidi.nl/en/employees/frans-willekens/

Comparative Population Studies

www.comparativepopulationstudies.de

ISSN: 1869-8980 (Print) – 1869-8999 (Internet)

Published by
Federal Institute for Population Research (BiB)
65180 Wiesbaden / Germany

E-mail: cpos@bib.bund.de