

International Journal of Population Data Science

Journal Website: www.ijpds.org



How Data Linkage Is Applied to Build Longitudinally Linked Data and Genealogy Data Based on Massive Sources of Administrative Data: A Two-decade Research on Linking Taiwan Household Registers

JI-PING LIN¹

¹Academia Sinica, Taipei, Taiwan

My presentation uses data linkage to build data on Taiwan indigenous peoples (TIPs) as an example. TIPs used to be invisible and marginalized in Taiwan. In recent years, TIPs not only become very visible in Taiwan, but are also known globally. The most important contribution that helps TIPs become visible and are empowered in the real world is successfully building a number of big open data sets (see TIPD at <https://osf.io/e4rvz/>), based on open science, open data, data science, and scientific computing. Central to the processes of building open data are data linkage, including deterministic and probabilistic, that integrate Taiwan household registers and other administrative data from 2007 to 2024. Using data linkage, I have successfully built longitudinally linked register big data of population dynamics, with individual-level spatial information (point) and temporal information (monthly) being integrated in the linked data. In my presentation, I will demonstrate data linkage in building big complexed liked

data (e.g., longitudinal and genealogy data), with a particular emphasis on the role of fine-tuning computing infrastructure (e.g., accelerating CPUs, DRAM, data transfer buses between CPUs-DRAM and between CPUs-PCIe lanes) to accelerate computing speed while conducting massive data linkage. In addition, I will demonstrate an automated geocoding method that allows us to parse spatial information from Google Map quickly and how legal and ethical issues are resolved. For reference, see <https://link.springer.com/article/10.1007/s43545-025-01049-1>.

Keywords

computing acceleration, data linkage, data science, longitudinally linked data, open science, scientific computing.

