

How Digital and Computational Demography Make Hard-to-reach Population Visible: Evidence from a Two-decade Research on the Population of Taiwan Indigenous Peoples

Ji-Ping Lin, Ph.D.

Academia Sinica

Nankang, Taipei

Taiwan 115

Email: jplin@sinica.edu.tw

ORCID: <https://orcid.org/0000-0003-2348-9243>

Long Abstract

The emerging availability of enriched big digital population data and advance in computing capability of digital infrastructure in the past two decades have not only helped overcome traditional constraints in population studies, but also are reshaping our world and transforming conventional thoughts about decision and policy making. These advances in the past two decade serve as the foundation of digital and computational demography. However, big digital population data does not essentially offer us enough insight and vision. We need to go further to extract useful information and improve data quality to fit real-world circumstances through the processes of enriching and integrating spatial and temporal information. Taking full advantage of big digital population data requires not only knowledge about fundamentals of computational demography and data science, but also the ability of implementation.

My research focuses on determinants of digital and computational demography, i.e., the role of scientific computing, data science, and open science in the context of extracting valuable information embedded in source individual data, enriching the extracted information through the processes of cleaning, cleansing, crunching, reorganizing, and reshaping the source data. The data enrichment processes produce a number of data sets that contain no individual information but retain most of the source data information. The enriched data sets thus can be open to the public as open data. Because the corresponding domain knowledge about hard-to-reach population research and Taiwan Indigenous Peoples (TIPs) is not easy to understand for the audience, my presentation will make a very short introduction about TIPs.

Central to my research is fundamentals of digital computational demography, geographical information science, data science, and open science. These research methods and techniques enable me to build a number of big open population data that help overcome long-term research data shortages in hard-to-reach population and ethical/legal issues. In addition, my research has developed a simple but fast geocoding method that allows us to parse spatial information from Google Map. Using record linkage method, 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 data.

My presentation aims at (1) addressing how open and enriched population data sets are built by integrating hacking skills, advanced math/statistics methods, and domain knowledge of various disciplines on the basis of data science and open science, and (2) the role of online open data repositories in promoting crowd collaboration. Main foci are as follows:

1. To demonstrate techniques in record linkage and highly precise address-matching geocoding that allow us to enrich temporal and spatial information in big data;
2. To highlight how data engineering and data sharing enable us to build and integrate open data repositories systematically and automatically;
3. To demonstrate the role of building open data in promoting online crowd collaboration and making hard-to-reach population visible to the real world, using the research on Taiwan indigenous peoples as an example.

Keywords: computational demography, data science, hard-to-reach population, open science, open data, Taiwan indigenous peoples

Reference:

1. Center for Open Science (COS). 2024. "Researcher Q&A: A Conversation About Research on Taiwan Indigenous Peoples and Making a Hard-to-Reach Population Visible", September 7. https://www.cos.io/blog/taiwan-indigenous-peoples?hs_preview=QUOrJLoo-176808682520
2. Lin, Ji-Ping. 2021. "Computational Archives of Population Dynamics and Migration Networks as a Gateway to Get Deep Insights into Hard-to-Reach Populations: Research on Taiwan Indigenous Peoples" *Proceedings of 2021 IEEE International Conference on Big Data*, IEEE Computer Society Press. DOI: <https://doi.org/10.1109/BigData52589.2021.9671838>
3. Lin, Ji-Ping. 2018. "Human Relationship and Kinship Analytics from Big Data Based on Data Science: A Research on Ethnic Marriage and Identity Using Taiwan Indigenous Peoples as Example" pp.268-302, in Stuetzer et al. (ed) *Computational Social Science in the Age of Big Data. Concepts, Methodologies, Tools, and Applications*. Herbert von Halem Verlag (Cologne, Germany), in Neue Schriften zur Online-Forschung of the German Society for Online Research (DGOF). <https://www.halem-verlag.de/produkt/computational-social-science-in-the-age-of-big-data>
4. Lin, Ji-Ping, 2017, "An Infrastructure and Application of Computational Archival Science to Enrich and Integrate Big Digital Archival Data: Using Taiwan Indigenous Peoples Open Research Data (TIPD) as Example" in *Proceedings of 2017 IEEE Big Data Conference*, the IEEE Computer Society Press. <https://doi.org/10.1109/BigData.2017.8258181>
5. Lin, Ji-Ping. 2017. "Data Science as a Foundation towards Open Data and Open Science: The Case of Taiwan Indigenous Peoples Open Research Data (TIPD)" in *Proceedings of 2017 International Symposium on Grids & Clouds, PoS (Proceedings*

of Science).

<https://newdoc.nccu.edu.tw/teasyllabus/1071254120001/DataScienceAsFoundationTowardOpenDataScienceISGC17PoSJPLin.pdf>.

Open population data:

1. Lin, Ji-Ping, Ming-Cheng Lee, Hiu Ha Chong, Li-Chuan Liu, Kui Kasirisir, and Hsin-Chung Wang. 2024. "TIPD : Taiwan Indigenous Peoples Open Research Data." OSF <https://osf.io/e4rvz/>. May 8. doi:10.17605/OSF.IO/E4RVZ.
2. Lin, Ji-Ping. 2024. "TPDD : Taiwan Indigenous Peoples Population Dynamics Open Data." OSF <https://osf.io/ukjgs/>. May 7. doi:10.17605/OSF.IO/UKJGS.
3. Lin, Ji-Ping. 2024. "TIHV : Taiwan Indigenous People's High-Resolution Visualization of Population Distribution, Migration Dynamics, Traditional Communities by Ethnic Groups." OSF <https://osf.io/v8zk3/>. May 8. doi:10.17605/OSF.IO/V8ZK3.
4. Integrated Query System of TICD: <https://TICDonGoogle.RCHSS.sinica.edu.tw>
5. TIPs Migration Dynamics: <https://www.rchss.sinica.edu.tw/capas/posts/11329>
6. High-resolution visualizations of population distribution, migration dynamics, traditional communities: <https://www.rchss.sinica.edu.tw/capas/posts/11393>
7. Interactive migration visualizations: https://www1.rchss.sinica.edu.tw/jplin/TIPD_Migration/

Appendix 1:

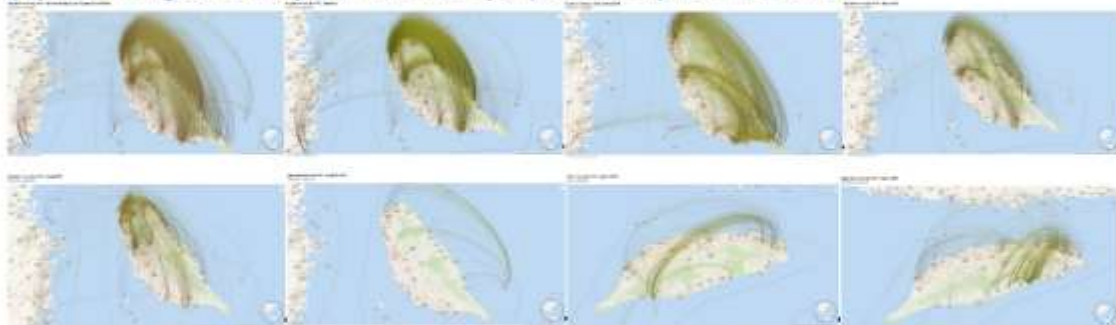
TIPD <https://www.rchss.sinica.edu.tw/capas/posts/11206>

TIPD: A Multidisciplinary Research in Open Science, Open Data, Open Source, Data Science, Big Data, and Scientific Computing (<https://TIPD.sinica.edu.tw> and <https://osf.io/e4rvz/>)

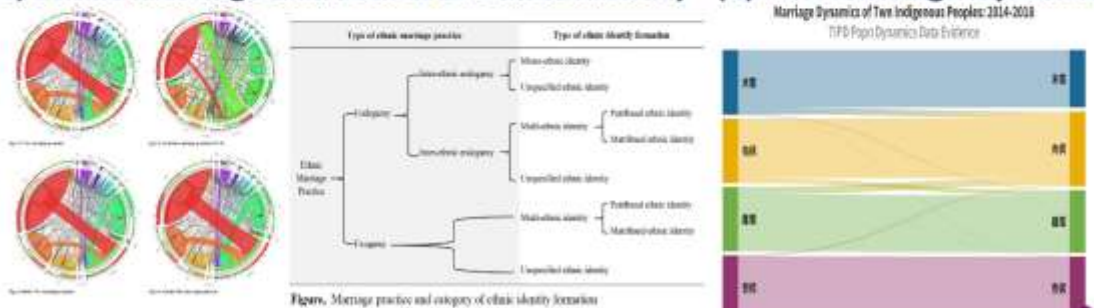
(1) Research Methodology (2) TIPs Distribution by Ethnic Groups(selected) (Dec. 2018)



(3) TIPs Migration by Ethnic Groups (selected): 2013-2018



(4) Ethnic Marriage and Source of Ethnic Identity (5) TIPs Marriage Dynamics



(6) TIPs Population Dynamics: 2013-2018



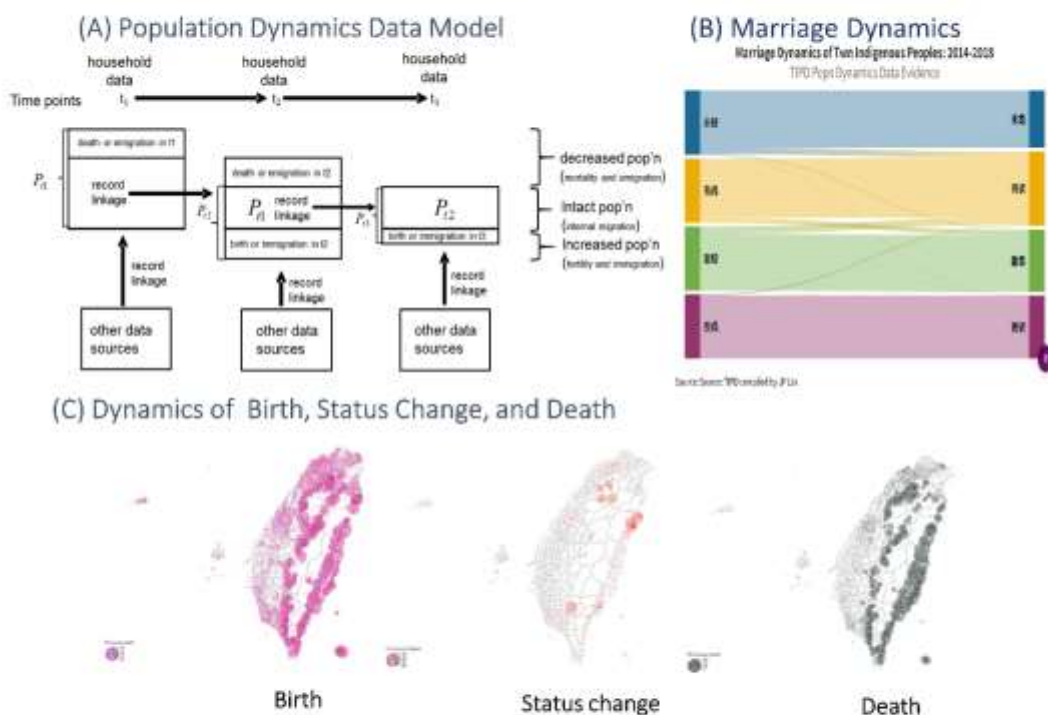
Credit: CopyLeft(L) 2019 TIPD joint research program

Appendix 2:

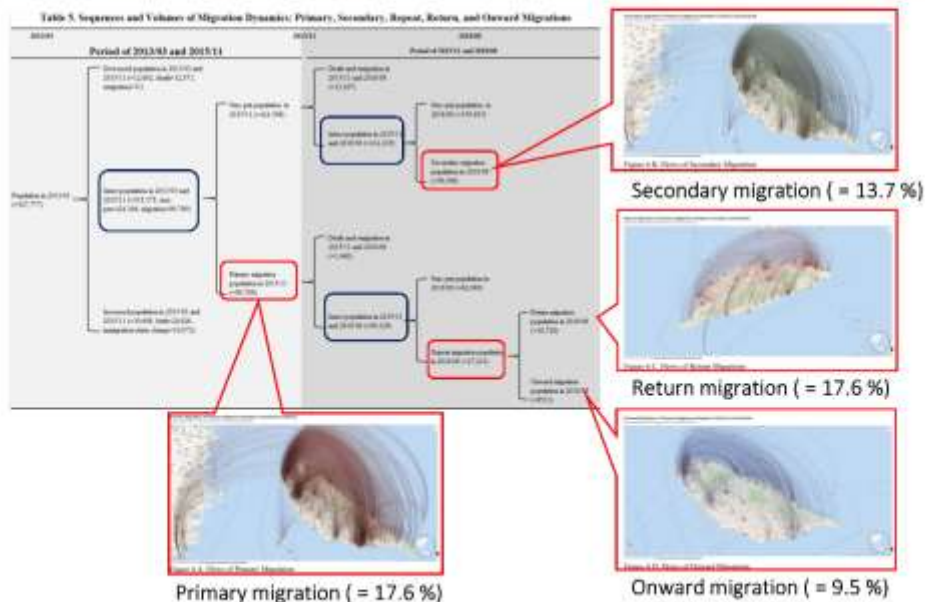
TPDD: <https://www.rchss.sinica.edu.tw/capas/posts/11621>

TPDD

Taiwan Indigenous Peoples Population Dynamics open Data
台灣原住民族人口動態開放資料 (<https://osf.io/ukjgs>)



(D) Migration Dynamics: Primary, 2nd-stage Primary, Return, and Onward Migrations



* TPDD viz built on 12 October, 2023

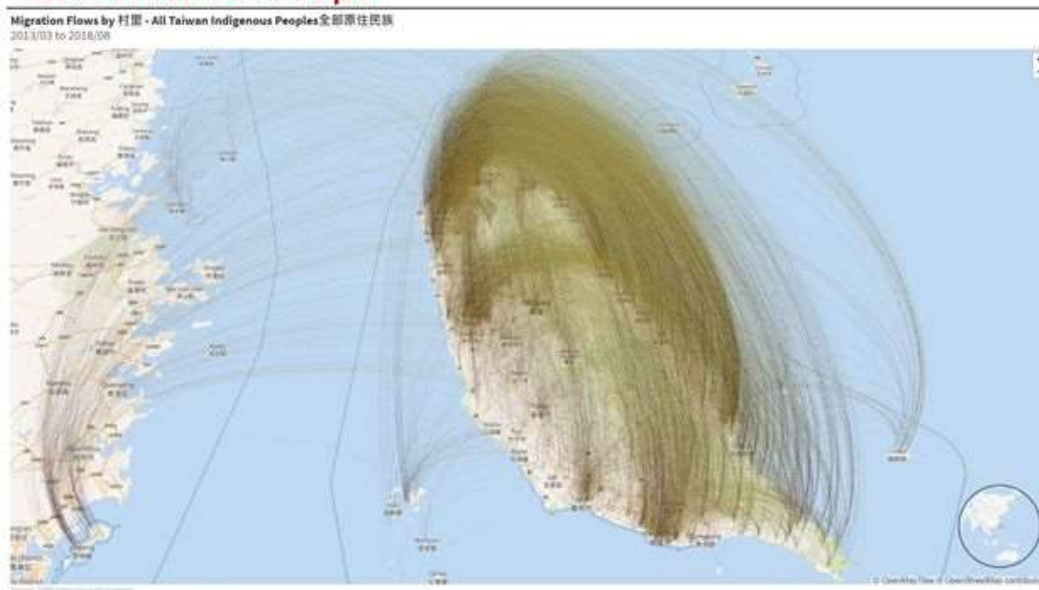
Appendix 3:

High-resolution visualizations of population distribution, migration dynamics, traditional communities: <https://www.rchss.sinica.edu.tw/capas/posts/11393>

台灣原住民人口分布、遷徙、傳統部落 10K高解析度圖



台灣原住民族人口遷徙：2013~2018 Population Migration of Taiwan Indigenous Peoples in 2013-2018 各族 All Ethnic Groups



Appendix 4:

TIPs Migration Dynamics: <https://www.rchss.sinica.edu.tw/capas/posts/11329>

Interactive migration visualizations:

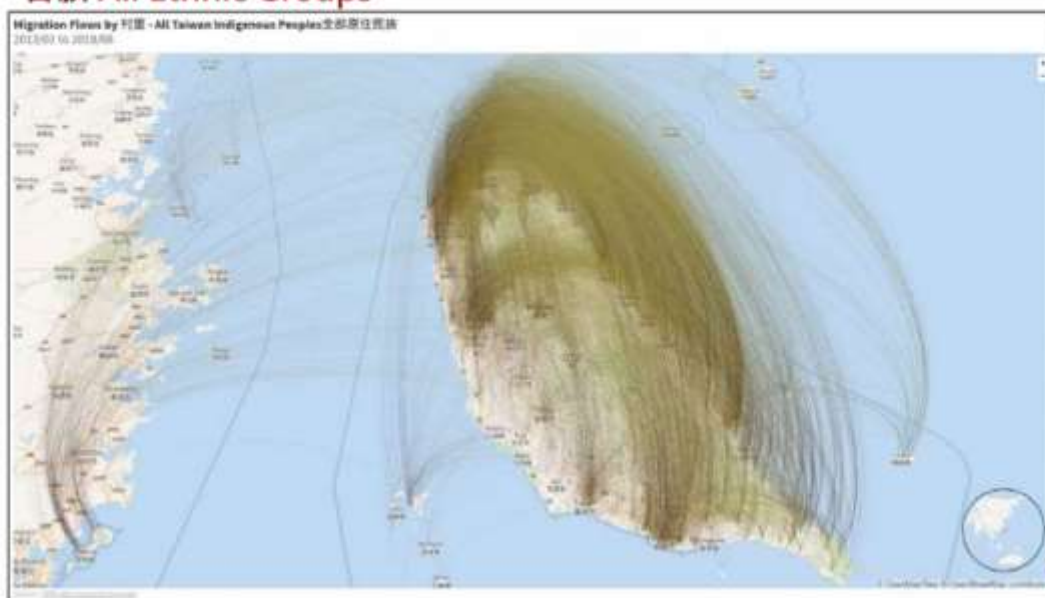
https://www1.rchss.sinica.edu.tw/jplin/TIPD_Migration/

台灣原住民遷徙動態(Migration Dynamics of TIPs)

台灣原住民族人口遷徙：2013~2018

Population Migration of Taiwan Indigenous Peoples in 2013-2018

各族 All Ethnic Groups



Source: 2013-2018 Population Dynamics, TIPD
Credit: TIPD research project (<https://posts.rchss.org/>)
Geocoding & Mapping: J.P. Lin, Academia Sinica
Published date: March 1, 2020

Primary Migration of Taiwan Indigenous Peoples in 2013/01 and 2018/08



Figure 6.A. Flows of Primary Migration

Primary migration

Return Migration of Taiwan Indigenous Peoples in 2013/01 and 2018/08



Figure 6.C. Flows of Return Migration

Return migration

Secondary Migration of Taiwan Indigenous Peoples in 2013/01 and 2018/08



Figure 6.B. Flows of Secondary Migration

Secondary migration

Onward Migration of Taiwan Indigenous Peoples in 2013/01 and 2018/08



Figure 6.D. Flows of Onward Migration

Onward migration

Appendix 5:

Integrated Query System of TICD: <https://TICDonGoogle.RCHSS.sinica.edu.tw>