

# Open Data and Open Science as Effective Ways to Revive Hard-to-reach Population: A Decade Research Applies to Taiwan Indigenous Peoples

Thursday, March 28, 2024 4:00 PM (30 minutes)

There has been a rich body of text and numerical archives about Taiwan indigenous peoples (TIPs) before 1940. Because detailed statistical archives and numerical data on TIPs were not available in the period between 1940-2010, we thus have very limited knowledge about developmental trajectory about TIPs. Such situation makes TIPs gradually become the so-called hard-to-reach population (HRP) and invisible in the real world. To overcome this issue, it becomes urgent to collect and build contemporary TIPs data. Supported by a decade research program from Council of Indigenous Peoples (<https://www.cip.gov.tw/>) in 2013, the author has been fully devoting himself to building a number of big open data on TIPS using scientific computing methods and techniques, based on the principles of open science and data science. The open data sets are built by integrating hacking skills, advanced math/statistics methods, and domain knowledge of various disciplines. Their repositories are hosted on OSF (Open Science Framework, <https://osf.io/>) and termed as TIPD (Taiwan Indigenous Peoples Open Research Data, for details, see <https://osf.io/e4rvz/>). TIPD complies with FARE data principle. It consists of the following categories of open data from 2013 to 2022: (1) categorical data, (2) multi-dimensional data, (3) population dynamics, (4) temporal geocoding data, (5) household structure data, (6) traditional TIPs community data (TICD), (7) generalized TICD query system, (8) genealogical data (not open to the public). The main contributions are as follows: (1) to enable TIPs who have been “invisible” to the world for seven decades to become “close” to “open” to the real world; (2) to empower TIPs researches from being “elite” to “ordinary” by using open data to reduce tech-barriers for researchers; (3) to extend TIPs studies from “local” to “global” arena by building bilingual open data repository; (4) to make research methods and techniques of TIPs switch from “macro” to “individual” level that make TIPS to revive from hard-to-reach to easy-to-reach population.

Keywords: Big Data, Data Science, FAIR, Open Data, Open Science, TIPD

## Reference

1. Utsurikawa, N., Miyamoto, N., Mabuchi, T. 1935. The Formosan Native Tribes: A Genealogical and Classificatory System. Institute of Ethnology, Taipei Imperial University (XXXXXXXXXXXXXXXXX, 1935,XXXXXXXXXXXXXXXXX,XXXXXXXXX).
2. Lin, Ji-Ping. 2017a. “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).
3. Lin, Ji-Ping, 2017b, “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.
4. 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 (Germany), Neue Schriften zur Online-Forschung of the German Society for Online Research.
5. 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.

## Reference of open data repositories:

6. TIPD: <https://www.rchss.sinica.edu.tw/capas/posts/11206>
7. TPDD: <https://www.rchss.sinica.edu.tw/capas/posts/11621>
8. TICD: <https://www.rchss.sinica.edu.tw/capas/posts/11205>
9. Integrated Query System of TICD: <https://TICDonGoogle.RCHSS.sinica.edu.tw>
10. TIPs Migration Dynamics: <https://www.rchss.sinica.edu.tw/capas/posts/11329>
11. High-resolution visualizations of population distribution, migration dynamics, traditional communities: <https://www.rchss.sinica.edu.tw/capas/posts/11393>
12. Interactive migration visualizations: [https://www1.rchss.sinica.edu.tw/jplin/TIPD\\_Migration/](https://www1.rchss.sinica.edu.tw/jplin/TIPD_Migration/)

**Primary author:** LIN, Ji-Ping (Academia Sinicca)

**Presenter:** LIN, Ji-Ping (Academia Sinicca)

**Session Classification:** Data Management & Big Data

**Track Classification:** Track 6: Data Management & Big Data