

International Journal of Population Data Science

Journal Website: www.ijpds.org



Leveraging Full Count Census Data through Record Linkage

Catherine A. Fitch¹, Victoria Udalova², and Luiza Antonie³

¹IPUMS, University of Minnesota

²U.S. Census Bureau

³The Canadian Peoples Project

Academic researchers in the U.S. and Canada have partnered with genealogical organizations and statistical agencies to create massive new scientific data collections from census enumerations. These data are creating new opportunities for research across the social and health sciences. By linking individuals and families across censuses, analysts can create national longitudinal panels that trace the characteristics of individuals over their lives and families over multiple generations.

IPUMS disseminates full count census enumerations for ten U.S. census years from 1850 to 1950. These full count data cover almost 800 million individual records and the IPUMS Multigenerational Longitudinal Panel (MLP) project links individuals' records across censuses. IPUMS data can be combined with data from the U.S. Census Bureau's data linkage infrastructure to link historical records to numerous recent censuses, surveys and administrative data that measure social, economic and health outcomes.

The Canadian Peoples (TCP) is a comprehensive public research database of 40 million coded and georeferenced records found in Canadian censuses from the middle of the nineteenth century until after the First World War. The data include personal, family and household characteristics of every individual enumerated in each census from 1852 to 1921. TCP investigators are exploring record linkage technology to solve challenges presented with these historical records.

The workshop included one presentation from each organization, with these objectives: introduce the data collections, explain the challenges and opportunities of census linkage, describe the linking strategies, and provide an overview on how to access the data.

