

# International Journal of Population Data Science

Journal Website: [www.ijpds.org](http://www.ijpds.org)



## Building the Longitudinal Education Outcomes Northern Ireland (LEO NI) Dataset: Linking Administrative Data for Policy Impact

Peter Wilgar<sup>1</sup>, Nichola McCullough<sup>1</sup>, Eoin Norris<sup>1</sup>, Brian Foley<sup>1</sup>, and Jana Ross<sup>1</sup>

<sup>1</sup>NISRA, Belfast, United Kingdom

### Objectives

This paper introduces the Longitudinal Education Outcomes Northern Ireland (LEO NI) dataset—an anonymised, research-ready resource that exemplifies the transformative potential of administrative data linkage. LEO NI integrates school-level records for individuals aged 14+ across eight academic years with seven years of post-16 data from Apprenticeships, Training Programmes, Further Education (FE), and Higher Education (HE). The dataset is designed to advance population data science by enabling empirical studies on educational attainment and post-16 trajectories, with direct implications for policy and practice.

translating methodological advances into actionable evidence, LEO NI demonstrates how linked administrative data can inform policy across sectors and jurisdictions, while addressing ethical and governance challenges.

### Methods

We detail methodological innovations in linking diverse administrative sources from the Department of Education and Department for the Economy, including School Census (with attendance), School Leavers Survey, and training/apprenticeship records. Linkage was achieved using anonymised identifiers under robust governance frameworks, ensuring privacy and compliance with legal standards. Additional variables capture Special Educational Needs (SEN) and health conditions, supporting complex, multi-dimensional analyses.

### Results

The resulting dataset comprises circa 0.5 million unique individuals and includes comprehensive metadata and a researcher guide. We outline quality assurance processes, linkage accuracy, and strategies for managing high-volume, heterogeneous data. These methodological insights contribute to best practices in large-scale data infrastructures. We highlight policy areas where this dataset can be used.

### Conclusion

Scheduled for release in Spring 2026 via secure research environments, LEO NI offers unprecedented opportunities for longitudinal analysis of education and employment outcomes. By

