

# Population Registers as Relational Databases

Tom Emery, Erasmus University Rotterdam,  
Jan van der Laan, Statistics Netherland, dj.vanderlaan@cbs.nl  
Marjolijn Das, Statistics Netherlands, m.das@cbs.nl  
Saskia te Riele, Statistics Netherlands, smm.teriele@cbs.nl  
Edwin de Jonge, Statistics Netherlands, e.dejonge@cbs.nl

## Abstract:

ODISSEI (Open Data Infrastructure for Social Science and Economic Innovations) is the national research infrastructure for the social sciences in the Netherlands. ODISSEI brings together researchers with the necessary data, expertise and resources to conduct ground-breaking research and embrace the computational turn in social enquiry. Through ODISSEI, researchers have access to large-scale, longitudinal data collections as well as innovative and diverse new forms of data. These can be linked to administrative data at Statistics Netherlands (CBS). Combining data from a wide range of sources enables researchers to answer new, exciting, interdisciplinary research questions and to investigate existing questions in novel, new ways. Here we present a whole population network. This uses administrative data to construct a relational administrative database, incorporating 1.4 billion relationships between the 17 million inhabitants of the Netherlands. The data is from the population registers in the Netherlands and includes robust persistent identifiers for linking individuals. The network is vast and was constructed using the ODISSEI Secure Super Computer which allows for the secure analysis of sensitive administrative data. Relationships are identified between individuals who live in the same household, live close to each other, work for the same company, attend the same educational institution, or belong to the same extended family. This produces a median number of 84 links per person for individuals in the population. This network has properties that are rare in observed social networks, which opens up new and exciting applications for network science in the social sciences. For more information about the data and the analytical opportunities it affords, please see the following pre-print: <https://osf.io/preprints/socarxiv/7jtb2/>