Traditional data sources (surveys, censuses)

Administrative sources with survey and other traditional data

Many possible types of data integration

New data sources with traditional data sources

Geospatial data with statistical information

Micro level data with data at the macro level

No generalised methodology exists; yet there are common solutions for shared problems

identifying needs, building partnerships, obtaining data, appropriate methods, managing quality, risks, comparability, metadata, standards, skills, …



StatsNZ Data integration manual, 2nd edition, lists the following as key steps in a data integration projects:

- establish need
- develop clearly defined objectives
- address legal, policy, privacy, and security issues
- consider provision of access to microdata and confidentiality of published outputs
- define governance structures and establish relationships with data providers and data users
- complete business case, privacy impact assessment, Government Statistician and agency chief executive sign-off, support from agencies supplying data
- obtain source data and gain a thorough understanding of data sources
- decide how to do the linking
- define and build information technology data storage and processing requirements
- carry out the linking
- validate the linking and provide quality measures
- open up access to linked data (if required)
- carry out the analysis and disseminate results.