DATA CONNECT AUSTRALIA

NCRIS National Research Infrastructure for Australia An Australian Government Initiative

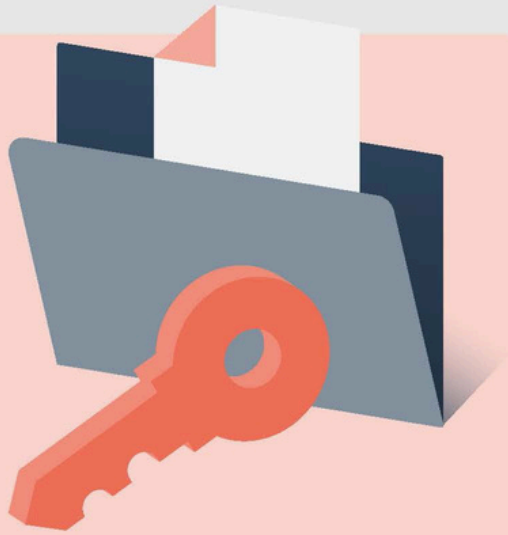
Protecting privacy

All **data linkage units** use privacy preserving methods to minimise the risk of people being reidentified in linked data. This means identifying information like names, dates of birth or addresses are removed from a dataset before it is provided to a researcher.

Why is privacy preservation important?

- Linked data research projects help provide new insights into the health of our population. However, patient confidentiality is crucial, which is why data linkage units take it so seriously. Your data linkage unit will advise you of any privacy concerns with your data request.
- Using linked data, researchers can confidently and ethically analyse a variety of datasets, gleaning important information from them, but they do not need to know the identity of each individual in the dataset. This is particularly important for projects where a waiver of consent has been granted by an independent human research ethics committee.

The separation principle



The separation principle is used to create separate and secure environments for data to be linked, stored, and analysed:

1 Data linkage units receive identifying data taken from a variety of datasets and linked records for the same person, using fields such as name, date of birth and address. These links, known as 'linkage keys', are stored separately to the identifying and health data attached to that record.

2 When a linked data project is approved, data linkage units extract the linkage keys from the datasets. Only linkage keys are required for this step, not identifying data.

3 The data required for analysis is attached to the linkage keys; no personal information needs to be reattached.

4 Researchers receive the data they need for their project and use the linkage keys to merge the datasets and conduct their analysis.