



2062.0

Information Paper

**Census Data
Enhancement Project: An
Update**

Australia

June 2006

New
Issue

Information Paper

**Census Data
Enhancement Project: An
Update**

Australia

June 2006

**Dennis Trewin
Australian Statistician**

AUSTRALIAN BUREAU OF STATISTICS

EMBARGO: 11.30AM (CANBERRA TIME) THURS 8 JUN 2006

ABS Catalogue No. 2062.0

ISBN 0 642 48221 7

© Commonwealth of Australia 2006

This work is copyright. Apart from any use as permitted under the *Copyright Act 1968*, no part may be reproduced by any process without prior written permission from the Commonwealth. Requests and inquiries concerning reproduction and rights in this publication should be addressed to The Manager, Intermediary Management, Australian Bureau of Statistics, Locked Bag 10, Belconnen ACT 2616, by telephone (02) 6252 6998, fax (02) 6252 7102, or email: <intermediary.management@abs.gov.au>.

In all cases the ABS must be acknowledged as the source when reproducing or quoting any part of an ABS publication or other product.

Produced by the Australian Bureau of Statistics

INQUIRIES

- For further information about these and related statistics, contact the National Information and Referral Service on 1300 135 070.

CONTENTS

page

CONTENTS

Preface	vii
Abbreviations	viii
Executive Summary	1

SECTION 1 INTRODUCTION

Introduction	3
Census data enhancement project	3
Overview of publication	3

SECTION 2 THE STATISTICAL LONGITUDINAL CENSUS DATASET

The Statistical Longitudinal Census Dataset	5
Benefits of the Statistical Longitudinal Census Dataset	5
Data involved in the Statistical Longitudinal Census Dataset	5

SECTION 3 QUALITY STUDIES

Quality studies	6
Census dress rehearsal to census matching	6
Quality study of Indigenous mortality data	6
Birth registration coverage quality study	6
Underenumeration of the Labour Force Survey	7
Linking DIMA's Settlement database and census data quality study	7
Assessing automatic matching for census post enumeration survey 2011	7
Linking the cancer register and census data quality study	8

SECTION 4 USE OF THE STATISTICAL LONGITUDINAL CENSUS DATASET WITH SPECIFIED NON-ABS DATASETS

Use of the Statistical Longitudinal Census Dataset with specified non-ABS datasets	9
Conditions of entry and settlement outcomes for migrants	9

SECTION 5 CONFIDENTIALITY AND PRIVACY

Confidentiality and privacy	10
Legislative protection	10
Destruction of census forms and name and address information	10
Access to ABS information	10
Data security	11
Meeting the obligations of the information privacy principles	11

ADDITIONAL INFORMATION

Glossary	13
----------	----

PREFACE

PREFACE

On 18 August 2005 I announced the ABS' intention to proceed with the Census Data Enhancement project to create the Statistical Longitudinal Census Dataset (SLCD) based on a 5% random sample. The modified proposal followed extensive consultations on the original proposal outlined in the Discussion Paper entitled *Enhancing the Population Census: Developing a Longitudinal View* (Cat. no. 2060.0) and the commissioning of an independent Privacy Impact Assessment (PIA).

This paper provides an update of the Census Data Enhancement project.

Dennis Trewin
Australian Statistician

ABBREVIATIONS

ABS	Australian Bureau of Statistics
DIMA	Australian Government Department of Immigration and Multicultural Affairs
LFS	Labour Force Survey
PES	post-enumeration survey
PIA	privacy impact assessment
SLCD	Statistical Longitudinal Census Dataset

EXECUTIVE SUMMARY

INTRODUCTION

The ABS has been set up, under the *Australian Bureau of Statistics Act 1975*, as the central statistical authority for the Australian Government. Among its functions, it is required to:

- collect, analyse and disseminate statistics and related information;
- avoid the duplication of collection, by official bodies, of information for statistical purposes; and
- achieve maximum possible utilisation, for statistical purposes, of information available to official bodies.

In undertaking its role, the ABS is required by the *Census and Statistics Act 1905* to maintain the confidentiality of information provided to it. ABS officers are required to sign an undertaking of fidelity and secrecy, and penalties are specified for any breach of this undertaking.

CENSUS DATA ENHANCEMENT PROJECT

On 26 April 2005, the ABS put forward a proposal to enhance the use of the Census of Population and Housing (*Enhancing the Population Census: Developing a Longitudinal View* Cat. no. 2060.0). This proposal included the creation of the Statistical Longitudinal Census Dataset (SLCD), and bringing together the SLCD with ABS and specified non-ABS datasets. This discussion paper was put forward for public comment.

In conjunction with the proposal, the ABS commissioned a Privacy Impact Assessment (PIA). The PIA said the proposal could operate consistently with the Privacy Act, assuming necessary attention by the ABS to informing individuals about the purpose and use of their information, and to data security. Section 5 shows that we have given necessary attention to these two points and therefore comply with the Privacy Act. Furthermore, in response to privacy concerns, the Census and Statistics Act has recently been amended to explicitly prevent individuals information from the census being subpoenaed for courts, tribunals or administrative purposes.

On 18 August 2005, after consideration of all comments received, the Australian Statistician announced the future of the project, via the *Census Data Enhancement - Statement of Intention*.

The project complies with the legislated function of the ABS to maximise the use, for statistical purposes, of information available to official bodies.

The potential benefits are substantial. They provide information on patterns in individual experiences over time and therefore provide insight into the effectiveness of policy or the need for new policy interventions. Examples of studies that could be undertaken are: pathways undertaken by migrants in their early years of settlement, particularly in employment; links between employment outcomes and education qualifications; transitions to higher education and work for young people from low income households; and the extent of income and employment mobility.

STATUS OF THE PROJECT

The project is currently investigating the feasibility of creating the SLCD. If feasible the SLCD will be:

- created by combining data from the 2006 Census with data from future censuses;
- created by bringing together the data through statistical techniques (see Glossary), rather than matching based on name and address; and
- based on a 5% random sample of the population.

EXECUTIVE SUMMARY *continued*

STATUS OF THE PROJECT *continued*

As part of the overall project and in line with the initial proposal and the statement of intention, the ABS is also undertaking several quality studies (see Glossary) which involve linking specified datasets to the 2006 Census data. These datasets are: death registrations; birth registrations; DIMA's Settlement database; Overseas Arrivals and Departures; the Cancer Register; 2005 Census Dress Rehearsal; Labour Force Survey (LFS); and the Census Post Enumeration Survey (PES). Only aggregate data will be disseminated from these quality studies.

More information about quality studies, and in particular the purpose of these studies is provided in Section 4 of the paper.

PRIVACY AND CONFIDENTIALITY PROTECTION

Respecting privacy and safeguarding confidentiality are fundamental tenets of ABS business, and have been core to its success over the 100 years of its history. The ABS has a very sound reputation in this regard. Both legislative protection and a very high security capability enable the ABS to ensure strong protection for this project.

The ABS is obligated to comply with provisions in the *Census and Statistics Act 1905* and the *Privacy Act 1988* to respect the privacy of individuals and to protect the confidentiality of their data. The provisions outlined under both these Acts will govern the use and release of data for the project.

In particular, the ABS will ensure the full protection of the secrecy provisions of the Census and Statistics Act apply to any dataset created through this project. This Act requires the ABS to keep information provided to it confidential. Identifiable data created through this project will not be provided outside the ABS, and only to those ABS staff who require it for their work.

DATA ACCESS AND AVAILABILITY

In accordance with the Census and Statistics Act, the use of non-identifiable unit record data (including the SLCD, and specified datasets brought together with the SLCD) would be available for statistical use, at the discretion of the Australian Statistician. This is outlined in Section 5 of this paper.

Access to potentially identifiable data will be restricted to ABS employees, with strict access controls. Where approved, researchers could only access non-identifiable unit records from the SLCD and this access would only be made through the ABS data laboratory, either by remote or on site access, both of which have strict security controls.

SECTION 1 - INTRODUCTION

INTRODUCTION

The ABS has been set up, under the *Australian Bureau of Statistics Act 1975*, as the central statistical authority for the Australian Government. Among its functions, it is required to:

- collect, analyse and disseminate statistics and related information;
- avoid the duplication of collection, by official bodies, of information for statistical purposes; and
- achieve maximum possible utilisation, for statistical purposes, of information available to official bodies.

In undertaking its role, the ABS is required by the *Census and Statistics Act 1905* to maintain the confidentiality of information provided to it. ABS officers are required to sign an undertaking of fidelity and secrecy, and penalties are specified for any breach of this undertaking.

CENSUS DATA ENHANCEMENT PROJECT

On 26 April 2005, the ABS put forward a proposal to enhance the use of the Census of Population and Housing (*Enhancing the Population Census: Developing a Longitudinal View* Cat. no. 2060.0). This proposal included the creation of the Statistical Longitudinal Census Dataset (SLCD), and bringing together the SLCD with ABS and specified non-ABS datasets. This discussion paper was put forward for public comment.

In conjunction with the proposal, the ABS commissioned a Privacy Impact Assessment (PIA). The PIA said the proposal could operate consistently with the Privacy Act, assuming necessary attention by the ABS to informing individuals about the purpose and use of their information, and to data security. Section 5 shows that we have given necessary attention to these two points and therefore comply with the Privacy Act. Furthermore, in response to privacy concerns, the Census and Statistics Act has recently been amended to explicitly prevent individuals information from the census being subpoenaed for courts, tribunals or administrative purposes.

On 18 August 2005, after consideration of all comments received, the Australian Statistician announced the future of the project, via the *Census Data Enhancement - Statement of Intention*.

The project complies with the legislated function of the ABS to maximise the use, for statistical purposes, of information available to official bodies.

The potential benefits are substantial. They provide information on patterns in individual experiences over time and therefore provide insight into the effectiveness of policy or the need for new policy interventions. Examples of studies that could be undertaken are: pathways undertaken by migrants in their early years of settlement, particularly in employment; links between employment outcomes and education qualifications; transitions to higher education and work for young people from low income households; and the extent of income and employment mobility.

OVERVIEW OF PUBLICATION

The ABS will periodically provide updates on work being undertaken as part of the Census Data Enhancement project. This paper provides the first update with respect to the project and the quality studies currently planned as part of it.

Section 2 discusses the SLCD and how it will be created.

SECTION 1 - INTRODUCTION *continued*

OVERVIEW OF PUBLICATION *continued*

Section 3 describes potential statistical studies using datasets based on the 5% sample used to create the SLCD, combined with specified non-ABS datasets using statistical techniques (see Glossary).

Section 4 describes the quality studies (see Glossary) being undertaken within the ABS.

Section 5 discusses confidentiality and privacy as it relates to the data held by the ABS.

SECTION 2 - THE STATISTICAL LONGITUDINAL CENSUS DATASET

THE STATISTICAL LONGITUDINAL CENSUS DATASET

The ABS is currently investigating the feasibility of creating the SLCD. If feasible, the SLCD will be:

- created by combining data from the 2006 Census with data from future censuses;
- created by bringing together the data through statistical techniques, rather than matching based on name and address; and
- based on a 5% random sample of the population.

BENEFITS OF THE STATISTICAL LONGITUDINAL CENSUS DATASET

The five-yearly census provides considerable information about Australian people and households.

The richness of the census data lies in the information it provides on topics such as:

- family structure;
- education and qualifications;
- work, including hours worked, occupation and industry;
- income; and
- housing.

Patterns in individual experiences over time in these areas, and factors that might influence these experiences, would be those where the SLCD would provide most insight. Examples include studies of the impact of parents' education and labour force participation on their children's subsequent participation, and the impact of factors such as family structures and household income on these outcomes.

The value of the SLCD will grow as data are brought together from future censuses. The SLCD by its nature is an investment in the future, and its value will be in answering research questions of the future.

DATA INVOLVED IN THE STATISTICAL LONGITUDINAL CENSUS DATASET

The creation of the SLCD itself will only involve the use of data from the Census of Population and Housing.

The SLCD will be based on a 5% random sample of persons from the 2006 Census.

Data collected from the 2011 Census will be combined with the data provided in the 2006 Census. Data for a sample of people covered in the 2011 Census but not covered in the 2006 Census will be added to the SLCD. This will continue for future censuses. The value of the SLCD will increase from 2011 on.

SECTION 3 - QUALITY STUDIES

QUALITY STUDIES

During the period of census processing, which for the 2006 Census is expected to be from August 2006 to November 2007, the ABS uses name and address information collected in the census to assist in processing the census data including the coding of family structure and checking undercount. After census processing is complete, all names and addresses held by the ABS are destroyed¹.

This section describes the studies currently planned by the ABS to use the 2006 Census data with other datasets at the time the census is being processed, using names and addresses to bring the datasets together. The resulting datasets will be destroyed at the completion of the studies.

CENSUS DRESS REHEARSAL TO CENSUS MATCHING

This study aims to test the feasibility of bringing together a 5% sample of one census with subsequent censuses, without using names and addresses, ie by statistical techniques. As a model for this, the 2005 Census Dress Rehearsal will be linked to the 2006 Census both with and without names and addresses as matching variables. Linking with names and addresses, while not perfect, will provide a benchmark for assessing linkage quality when statistical techniques are used. Birth and death registration information, as well as overseas arrivals and departures, may be used to resolve cases of unlinked records.

QUALITY STUDY OF INDIGENOUS MORTALITY DATA

Death registrations are provided to the ABS by the State and Territory Registrars of Births, Deaths and Marriages. Current estimates indicate that, while most deaths are registered, only approximately 60% of Indigenous deaths are identified as being Indigenous. The remainder have Indigenous status recorded as not stated or not Indigenous. This rate of undercoverage substantially impacts on the quality of the mortality data, including cause of death data. This impedes the ability to monitor trends in health status and particular causes of mortality. The study therefore has strong relevance to developing a better understanding of, and improving, the health status and well-being of the Indigenous population.

This study would link data from the 2006 Census to death registrations for August 2006 to August 2007 to assess the match between Indigenous status in the two collections. The expected outcomes include: gaining an understanding of differences in Indigenous status recording between death registration and census data; identifying factors contributing to undercoverage of Indigenous deaths; and investigating the feasibility of calculating and applying adjustment factors for Indigenous deaths output data.

BIRTH REGISTRATION COVERAGE QUALITY STUDY

Birth registrations are provided to the ABS by the State and Territory Registrars of Births, Deaths and Marriages. While there is a legislative requirement to register a birth, generally within 60 days, this does not always happen. There may be registration lags and processing lags. The ABS is usually able to take account of processing lags in compiling the data, but the registration lags are much harder to estimate.

¹ For the 2001 and future Censuses, all people who complete the census form are provided with the option of having their name identified census information stored for the use of future generations. The information for those people who elect to have their information retained is in the custody of the National Archives of Australia, and is not accessible to the ABS, or any other organisation or person for 99 years.

SECTION 3 - QUALITY STUDIES *continued*

BIRTH REGISTRATION COVERAGE QUALITY STUDY *continued*

The aim of the study is to assess the quality of birth registration data in regard to its coverage and delays in registration, and to derive factors that might be applied to raw birth registration data for the 2006 - 2011 intercensal period. To do this, birth registrations for the preceding three to four years will be linked with records for young children in the 2006 Census. The linked dataset would also enable the ABS to undertake quality assessment of fertility estimates, improve the quality of preliminary population estimates and provide information to State and Territory registrars about patterns in the characteristics of people who register births late or not at all.

UNDERENUMERATION OF THE LABOUR FORCE SURVEY

The Labour Force Survey (LFS) provides estimates of employment and unemployment used to set key economic and social policy parameters. The LFS has an extensive quality assurance program aimed at ensuring the ABS provides users with labour force estimates of the highest quality. A key aspect of this program is minimising non-sampling error. Non-sampling error occurs in all surveys. It arises from inaccuracies in collecting, recording and processing the data. Non-sampling error also arises because it is not always possible to obtain information for all persons selected in the survey. Underenumeration is caused because some groups within the population, for example young males, can be more difficult to contact than others. If this occurs, the characteristics of the group being underenumerated are potentially not fully reflected in the estimates from that survey.

An investigation into underenumeration in the LFS is a component of the LFS quality assurance program. The 2006 Census provides an opportunity to bring information related to dwellings in both the LFS and the 2006 Census together. The objective of the study is to gain a better understanding of the characteristics of the people who are missed in the LFS. This will assist in the development of procedures to minimise underenumeration in the future.

LINKING DIMA'S SETTLEMENT DATABASE AND CENSUS DATA QUALITY STUDY

The aim of the study is to test the feasibility of bringing together the SLCD with DIMA's Settlement database using statistical techniques. The 2006 Census processing period provides the opportunity to attempt to link the datasets using two approaches, both with and without names and addresses as matching variables. Linking with names and addresses, while not perfect, will provide a benchmark for assessing linkage quality when statistical techniques are used. This information will be used to determine if the datasets produced for the statistical study outlined in Section 4 will be fit for their purpose, and aid interpretation of the results of the statistical analysis.

ASSESSING AUTOMATIC MATCHING FOR CENSUS POST ENUMERATION SURVEY 2011

The PES is conducted a few weeks after the Census of Population and Housing to estimate the number of people who were missed in the census or who were counted more than once. The PES uses an independent sample of people and collects their name, key demographic items, their address on census night and other addresses where they may have been counted ('search addresses'). The name and demographic items are used to match a person to the census returns using the search addresses. The process of matching persons is manual and therefore is slow and expensive. Because of the cost, person matching is restricted to the search addresses, while the ideal is to find whether a person was counted at any address in the census, particularly if the address supplied is imprecise.

SECTION 3 - QUALITY STUDIES *continued*

ASSESSING AUTOMATIC MATCHING FOR CENSUS POST ENUMERATION SURVEY 2011 *continued*

The aim of the study is to improve the quality and efficiency of the matching processes of PES by assessing the feasibility and benefits of largely automating this process. Automation also gives the opportunity to 'widen the search area' when matching, so that more actual matches are detected. Automation would not eliminate clerical judgement completely but would output borderline cases for clerical resolution.

The accuracy of the PES matching undertaken each census is key to the quality of Australian and State level population estimates.

LINKING THE CANCER REGISTER AND CENSUS DATA QUALITY STUDY

A quality study relating to cancer registrations, is being considered. If it proceeds, this study would aim to test the feasibility of bringing together the SLCD with the Cancer Register. At this stage there are issues associated with limits to the address information on the Cancer Register, the currency of the data and the availability of ABS resources. While the aim of this study would be to enhance the SLCD for research so that the past and current characteristics of people diagnosed with cancer can be analysed, it is not yet clear whether this type of analysis could be supported by the SLCD.

SECTION 4 - USE OF THE STATISTICAL LONGITUDINAL CENSUS DATASET WITH SPECIFIED NON-ABS DATASETS

USE OF THE STATISTICAL LONGITUDINAL CENSUS DATASET WITH SPECIFIED NON-ABS DATASETS

The SLCD will be used further by bringing it together with specified non-ABS datasets using statistical techniques. The *Census Data Enhancement - Statement of Intention* indicated that the use of the 5% SLCD with specified non-ABS datasets will proceed for particular datasets, only with the agreement of the custodians of the non-ABS datasets and only for statistical purposes.

The specified non-ABS datasets are:

- birth and death register data, including cause of death;
- DIMA's Settlement database (see Glossary); and
- national disease registers (see Glossary).

At this early stage only one such statistical study is planned and it is discussed below. This, and other statistical studies will be contingent on the quality studies described in Section 3. The outcomes of the quality studies will indicate if valid statistical analyses can be performed, and the types of statistical study for which the SLCD in combination with the above datasets, is most useful.

CONDITIONS OF ENTRY AND SETTLEMENT OUTCOMES FOR MIGRANTS

On arrival in Australia, information is collected in relation to permanent migrants, including the permanent visa category under which they arrive (family, humanitarian or skilled). This information, statistically matched to the census, will provide valuable insights into patterns of settlement of the different groups of migrants, including family formation, housing, labour force characteristics, changing occupations, educational pathways and region of settlement. This information will help government assess immigration policies to ensure they deliver maximum benefits to the Australian economy and society. It will also assist in the service delivery to support the effective settlement of permanent migrants to Australia.

This statistical study is dependent upon the results of the related quality study (Linking DIMA's Settlement database and census data quality study) outlined in Section 3.

SECTION 5 - CONFIDENTIALITY AND PRIVACY

CONFIDENTIALITY AND PRIVACY

A fundamental aspect of the Census Data Enhancement project is the management of privacy. The ABS is obligated to comply with provisions in the *Census and Statistics Act 1905* and the *Privacy Act 1988* to respect the privacy of individuals and to protect the confidentiality of their data. The provisions outlined under both these Acts will govern the use and release of data from this project.

This section describes the processes that the ABS has in place to meet its obligations under both pieces of legislation.

LEGISLATIVE PROTECTION

The *Census and Statistics Act 1905* protects the confidentiality of data provided to the ABS. These protections apply to all data collected by, or supplied to the ABS, including the data to be used for this project.

In particular, the ABS will ensure the full protection of the secrecy provisions of the Census and Statistics Act applies to any dataset created through this project. This Act requires the ABS to keep information provided to it confidential. Identifiable data created through this project cannot, by law, be provided outside of the ABS. To strengthen these provisions, the Census and Statistics Act has recently been amended to explicitly prevent individuals information from the census being subpoenaed for courts, tribunals or administrative purposes.

All ABS officers are legally bound to secrecy under the Census and Statistics Act. Officers sign an undertaking of fidelity and secrecy to ensure that they are aware of their responsibilities. Section 19 of the Act forbids past or present ABS officers from directly or indirectly divulging information collected under this Act, by imposing penalties of fines of up to \$13,200 or imprisonment for two years, or both, for any breaches.

DESTRUCTION OF CENSUS FORMS AND NAME AND ADDRESS INFORMATION

The ABS maintains a long-standing practice of destroying the individual census forms once processing of the census information is complete. Names and addresses collected in the census are not retained by the ABS, after completion of census processing.

For the 2001 and future Censuses, all people who complete the census form are provided with the option of having their name identified census information stored for the use of future generations. The information for those people who elect to have their information retained is in the custody of the National Archives of Australia, and is not accessible to the ABS, or any other organisation or person for 99 years.

ACCESS TO ABS INFORMATION

Procedures are put in place to ensure all aggregate outputs disseminated by the ABS are not likely to enable the identification of an individual or household.

The Census and Statistics Act allows the ABS to make non-identifiable unit record data available for statistical use, at the discretion of the Australian Statistician. When this occurs, tight procedures and protocols govern such access. The ABS scrutinises the data, closely manages the dissemination channels and requires legal undertakings from individuals and their organisations to ensure that providing access to data is not likely to enable the identification of an individual. All users who request access to these non-identifiable datasets must state their intended statistical purpose for using the data. Should a user breach the conditions of the undertaking and deliberately attempt to identify an individual, they are subject to prosecution under the Census and Statistics Act.

SECTION 5 - CONFIDENTIALITY AND PRIVACY *continued*

ACCESS TO ABS INFORMATION *continued*

Access to non-identifiable unit record data from the SLCD, and the use of the SLCD in conjunction with specified non-ABS data sources, will be subject to all the above procedures and will be restricted to access through the ABS data laboratory, either by remote or on site access.

DATA SECURITY

The ABS maintains practices of a high standard to ensure the security of all information it holds. Features of the ABS environment are:

- strong security arrangements for all ABS information technology systems. ABS conforms with IT Security arrangements set out in the Australian Government Information Security Manual ASCI 33;
- strict control of access to all ABS premises in accordance with the Commonwealth Protective Security Manual to ensure compliance with legislative responsibilities;
- appropriate personnel security arrangements. Upon appointment, all ABS staff undergo security checks and are required to sign an undertaking of fidelity and secrecy;
- a secured Internet gateway which is reviewed annually by Defence Signals Directorate;
- regular Protective Security risk reviews to ensure that security arrangements continue to be effective; and
- an ongoing program of security audits and reviews of computer systems and the physical environment.

The ABS induction and training strategy for its staff places strong emphasis on the importance of security in safeguarding confidentiality, and on the appropriate use of the technology environment.

In addition, to ensure the security of the data being used in the Census Data Enhancement project, all data will be stored in a separate server at the Census Data Processing Centre. Access to the server will be restricted to those who need access. Once census processing has been completed, name and address identified data on the Census Data Enhancement server will be destroyed.

Further details about data security are available on the ABS website at www.abs.gov.au.

MEETING THE OBLIGATIONS OF THE INFORMATION PRIVACY PRINCIPLES

The PIA said the proposal could operate consistently with the Privacy Act, assuming necessary attention by the ABS to informing individuals about the purpose and use of their information, and to data security. The previous subsection discusses the arrangements put in place to ensure security. The need to notify individuals is addressed in the Census Information Booklet (referred to as the Household Guide). This is provided to all households with the census form.

A Census Information Booklet is provided to all households with their census form. At the very beginning of the guide there will be statements on whether completing the census form is compulsory, privacy and confidentiality, and adding value to the census. The latter deals with the Census Data Enhancement project and is very prominent in the guide. The actual statement is shown below.

SECTION 5 - CONFIDENTIALITY AND PRIVACY *continued*

MEETING THE OBLIGATIONS OF THE INFORMATION PRIVACY PRINCIPLES *continued*

"For the 2006 Census the ABS will be developing a new way of looking at census data that will add value to it by combining information from a sample of individual records over time between successive censuses. This information will also be used with information on specified events such as birth and death registrations, cancer registrations, and permanent settlement and departure information, for that sample.

This will help to paint a picture of how different situations impact on outcomes for people by expanding the range of statistical information available from the census. It will help answer questions such as 'why are people where they are today?', 'how did they get to where they are?' and 'what can be done to improve people's situation?'

The sample of records will be brought together for successive censuses and events using variables such as age or date of birth, sex, country of birth and geographic area. This will not involve the ABS keeping name and address information for the census, and by law no data that is likely to enable the identification of a person can be released from the ABS.

For more information on this, go to www.abs.gov.au or call 1300 362 883."

A Census Data Enhancement Fact Sheet will also be produced. This will provide additional information for those who visit the web site or call the Census Inquiry Service for additional information.

GLOSSARY

ABS data laboratory	The ABS provides a service for use of unit record data through an access system involving relevant analytical software. Statistical analyses specified by researchers are submitted and run against the dataset to produce statistical outputs. The system prohibits certain types of analyses that would have the potential to provide identifiable data, while allowing most standard analytical tools to be used against the unit data. Non-identifiable outputs of the analyses are provided to the researcher. The unit record data remains within the secure ABS environment. This relates to both the remote access and on site laboratories.
Birth and death registrations	Data obtained from the registration of births and deaths. The registration of births and deaths is the responsibility of the State and Territory Registrars of Births, Deaths and Marriages. As part of the death registration process, information as to the cause of death is supplied by the medical practitioner certifying the death or by a coroner. Other information about the deceased is supplied by a relative or other person acquainted with the deceased, or by an official of the institution where the death occurred. For births, the data are based on the information form completed by the parents of the child.
Census processing period	The period of time immediately after the conduct of the Census of Population and Housing during which the census forms are processed to produce statistical outputs. For the 2006 Census, this period is expected to extend from August 2006 to November 2007.
Dataset	A file containing the individual responses from a statistical collection, administrative records or register of information (for example disease register). Datasets are used to generate statistical output.
DIMA's Settlement database	Statistical data held by the Department of Immigration and Multicultural Affairs (DIMA) collected from permanent visa applications and other supplementary information sources. Includes information on visas granted for permanent entry into Australia from offshore and for permanent resident status onshore. It does not include data on non-visa settlers (such as New Zealand citizens).
Identifiable	In this publication, unit record data is considered identifiable if it is possible to use the data available in the record to identify the specific individual to whom it refers.
Longitudinal dataset	A dataset which contains information collected for the same person or unit over a number of different points in time.
National disease registers	Registers containing information about people diagnosed with certain diseases have been established in the states and territories. These registers allow research about the particular disease. To ensure a coordinated approach to the use of these registers, the Australian Institute of Health and Welfare (AIHW) administer national disease registers, consolidating the information contained in the state and territory registers.
Quality studies	In this publication, quality studies refer to the datasets that have been brought together using name and address information, for ABS internal studies that look at the quality of the data. These studies can look at the quality of linkages when different linking techniques are used, or they can look at the quality of data that the ABS are currently disseminating, when compared with the information contained in the census dataset.
Statistical purposes	Functions related to the compilation, analysis and dissemination of statistics. Statistical purposes precludes use of a dataset for administrative or client management purposes, where there is an impact on specified individuals.
Statistical techniques	In this publication, statistical techniques refer to the method that would be used to bring together census data over time to form the SLCD, and to bring the SLCD together with specified non-ABS datasets. The proposed method is often referred to as probabilistic record linkage, which involves bringing together data from two different datasets using a number of characteristics such as age/date of birth, sex, geographic region, and country of birth. All possible linkages based on these data items are evaluated. The records for which the linkage is most likely to be correct are brought together. The statistical

GLOSSARY *continued*

Statistical techniques techniques used do not involve the use of name and address information to bring
continued together the datasets.

FOR MORE INFORMATION . . .

INTERNET **www.abs.gov.au** the ABS web site is the best place for data from our publications and information about the ABS.

LIBRARY A range of ABS publications are available from public and tertiary libraries Australia wide. Contact your nearest library to determine whether it has the ABS statistics you require, or visit our web site for a list of libraries.

INFORMATION AND REFERRAL SERVICE

Our consultants can help you access the full range of information published by the ABS that is available free of charge from our web site, or purchase a hard copy publication. Information tailored to your needs can also be requested as a 'user pays' service. Specialists are on hand to help you with analytical or methodological advice.

PHONE 1300 135 070

EMAIL client.services@abs.gov.au

FAX 1300 135 211

POST Client Services, ABS, GPO Box 796, Sydney NSW 2001

FREE ACCESS TO STATISTICS

All ABS statistics can be downloaded free of charge from the ABS web site.

WEB ADDRESS **www.abs.gov.au**



2206200005069
ISBN 0 642 48221 7

RRP \$11.00