



## Economic and Social Council

Distr.: General  
16 April 2013

English only

---

### Economic Commission for Europe

#### Conference of European Statisticians

##### Sixty-first plenary session

Geneva, 10-12 June 2013

Item 4 (a) of the provisional agenda

##### Drivers for micro-data access

### **Australian Commonwealth government data integration initiatives**

#### **Note by the Australian Bureau of Statistics**

##### *Summary*

The paper discusses the key drivers for the development of an Australian Statistical Data Integration Framework for the Commonwealth Government. The paper also covers the framework, the results of the work so far and some of the challenges and work still ahead.

Demand for integration of large datasets at a unit record level is increasing in Australia. They are needed especially to assist in researching complex problems that have multiple causes and cut across many areas of government responsibility, such as climate change, joblessness, homelessness, obesity etc. While data integration is highly beneficial, the personal information of individuals has to be protected.

## I. Key drivers

1. There is an increasing demand for integration of large datasets at a unit record level in Australia. This demand is being driven by the potential of data integration to produce rich and dynamic datasets that can assist in researching complex ‘wicked problems’<sup>1</sup>, such as climate change, joblessness, homelessness and obesity. These issues often have multiple contributing and confounding causes, covering many areas of government responsibility, and therefore cannot be adequately researched and captured by one survey or administrative dataset.
2. The need to produce such rich and complex datasets in a low cost and timely manner has also driven the demand for data integration. Data integration harnesses the ‘collect once, use many times’ principle leading to efficient use of resources, lower respondent burden and the ability to produce evidence based policy quickly.
3. While the benefits from data integration may be great, there are demands on Australian government departments to protect the personal information of individuals, as set out in Australian privacy legislation governing all Commonwealth agencies. There is also a need to protect the confidentiality of data provided by individuals and individual businesses. This is because any failure by government to respect an individual’s or business’s right to privacy, or inadvertent breach of this right through poor management of data, has the potential to cause a high level of public concern.
4. If a major data breach occurred in relation to a sensitive integrated data set, it is likely that there would be a high level of public concern about, and a deleterious impact on public trust in, how all government departments treat personal and business data. This could in turn have an effect on the public’s willingness to provide information to government. The possible loss of trust in government is of particular concern for the Australian Bureau of Statistics (ABS), as the trust of providers is fundamental to being able to achieve high response rates to surveys and deliver accurate, high quality statistics.
5. The demand for data integration, coupled with the demands to maintain the privacy and confidentiality of those providing the data, prompted the Australian Statistician and the Secretary of the Department of Health and Aging (DoHA) to seek a Commonwealth approach to data integration for statistical and research purposes. This approach has led to all Commonwealth agencies working together towards a statistical data integration framework that aims to maximise the value of the data available in Commonwealth agencies, while managing the systemic risk to Australia’s statistical system if a breach occurs and public mistrust develops. These aims align with the Australian Government’s Gov2.0 initiative<sup>2</sup> which encourages co-operative leadership across government to realise the full social and economic value of public sector information as a national resource. It also aligns with the Blueprint for Reform of Australian Government Administration that notes the need for the Australian Public Service *‘to strengthen its capacity to undertake*

---

<sup>1</sup> Australian Public Service Commission, Tackling wicked problems: A public policy perspective (2007): [www.apsc.gov.au/publications-and-media/archive/publications-archive/tackling-wicked-problems](http://www.apsc.gov.au/publications-and-media/archive/publications-archive/tackling-wicked-problems)

<sup>2</sup> Government 2.0 Task Force, Engage: Getting on the Government 2.0 (December 2009): [www.finance.gov.au/publications/gov20taskforcereport/index.html](http://www.finance.gov.au/publications/gov20taskforcereport/index.html)

*rigorous research, gather and analyse data and provide the highest quality strategic policy advice.*<sup>3</sup>

6. The approach taken currently only involves Commonwealth government agencies. In Australia there are two other levels of government; state and territory government, and local government. Although the approach currently only covers Commonwealth government, it is hoped that in the future the concepts that lie behind this framework will be adopted by the other tiers of government (see the Future Directions section for more information).

## II. International work being conducted

7. The same drivers for data integration are fuelling the development of frameworks and guidelines in different countries all around the world. So far each country has taken a slightly different tack on the way integration of government data should be managed. Some countries, such as New Zealand and Canada, have taken a centralised approach where the national statistical organisations for each country conduct data integration for statistical and research purposes. They may conduct these projects on behalf of other government departments if requested and approved. Others, such as Scotland, have developed country wide principles on data integration and have put in place an oversight structure to recommend best practice and oversee data integration activities across the country. The principles are not binding and the responsibility of maintaining confidentiality of the data rests with the organisation integrating the data.

8. One of the most comprehensive frameworks the ABS is currently aware of is the framework proposed by the United Kingdom (UK) Administrative Data Task Force<sup>4</sup>, designed to put in place robust practices and procedures to minimise the risk of privacy and confidentiality breaches involving UK government data. The Task Force has recommended that dedicated national centres should be established to commission and undertake data linkage of government data, however, centre staff and researchers must not have access to personal identifiers and trusted third parties will be required to perform the linking. In addition, the report recommended these national centres develop and maintain information systems that track and record activity relating to data access and linking. Although this proposal has not yet been adopted, it demonstrates the need to move towards a whole of government approach to manage the privacy and confidentiality risks associated with data integration.

## III. Australian developments

9. Australia, like the UK's proposal, has focused on a whole of government approach to data integration and is currently testing a framework to integrate data safely for statistical and research purposes. The framework aims to reassure the public that their data, held by Commonwealth agencies, is being managed safely. It does this in two ways: firstly, by providing sufficient safeguards to minimise any potential risk to privacy and

<sup>3</sup> Advisory Group on Reform of Australian Government Administration, Ahead of the Game: Blueprint for Reform of Australian Government Administration, (March 2010): [www.dpmc.gov.au/publications/aga\\_reform/aga\\_reform\\_blueprint/index.cfm#blueprint](http://www.dpmc.gov.au/publications/aga_reform/aga_reform_blueprint/index.cfm#blueprint)

<sup>4</sup> The Administrative Data Task Force, The UK Administrative Data Research Network: Improving Access for Research and Policy, (December 2012): [www.esrc.ac.uk/funding-and-guidance/collaboration/collaborative-initiatives/Administrative-Data-Taskforce.aspx](http://www.esrc.ac.uk/funding-and-guidance/collaboration/collaborative-initiatives/Administrative-Data-Taskforce.aspx)

confidentiality; and secondly, by ensuring that statistical data integration is conducted in an open and accountable way.

## A. The process taken to develop the framework

10. The importance of establishing a Commonwealth framework for statistical integration was presented and endorsed in April 2009 by Portfolio Secretaries (a body later replaced in 2010 by the Secretaries Board, consisting of the heads of all Commonwealth government agencies and the Australian Public Service Commission). At this meeting Portfolio Secretaries also approved the development of a Cross Portfolio Statistical Integration Committee (CPSIC), jointly chaired by the Department of Health and Ageing and the Australian Bureau of Statistics, to pursue this work and report back on progress.

11. In February 2010 CPSIC reported back to Portfolio Secretaries seeking endorsement of a set of high level principles for data integration involving Commonwealth data for statistical and research purposes. The board endorsed the principles, which outline when data integration should occur, who is responsible for the data and the data integration process, how the data should be treated and the importance of transparency in the process of data integration (for the principles see Annex A).

12. Once agreement on the high level principles was secured, work began on implementing governance and institutional arrangements to support the application of the principles. An interim set of arrangements were agreed upon by the Secretaries Board in October 2010. The arrangements were considered as interim because of the desire to review the arrangements following an assessment of Commonwealth agency experiences before putting in place revised final arrangements. The review of agency experiences and the revised final arrangements are scheduled for completion by October 2013.

## B. What are the interim governance and institutional arrangements?

13. The main interim arrangements that have been established consist of:

(a) **A Cross Portfolio Data Integration Oversight Board.** The board is chaired by the Australian Statistician, and includes three to four departmental secretaries. The Oversight Board meets around four times per year and has been tasked with providing strategic and collaborative leadership for the integration of Commonwealth data. It assists in the management of the systemic risk associated with conducting multiple data integration projects and can endorse any changes or additions to the overall data integration environment;

(b) **The establishment of Integrating Authorities (IAs).** It was decided that there must be an IA for each statistical data integration proposal. The IA is responsible for the sound conduct of the data integration project and has overall responsibility to ensure that risks have been assessed, managed and mitigated throughout the duration of the project;

(c) **The establishment of Accredited Integrating Authorities.** It was also decided that an accreditation process should be in place for organisations that want to conduct 'high risk' data integration projects (such as very large complex projects, or ones integrating sensitive personal data). Once an organisation has met the requirements set out in the accreditation process, including independent auditing, the Oversight Board reviews the accreditation documentation and considers whether accreditation should be granted. Currently there are two accredited IAs, the Australian Institute of Health and Welfare

(AIHW) and the Australian Bureau of Statistics (ABS). However, more organisations intend to apply;

(d) **The development of the Best Practice Guidelines.** The guidelines will describe best practice for data integration involving Commonwealth data for statistical and research purposes. The guidelines are currently being drafted and are scheduled to be published by October 2013 when the final arrangements are put in place;

(e) **The development of a web-based register of projects.** Consistent with more open government, and building public trust through transparency, a web-based register has been established to include a description of all data integration projects involving Commonwealth data. Each IA conducting a project is responsible for adding their project/s to the register. A public feedback mechanism has also been included on the register to enable members of the public to indicate their support or concerns and to make suggestions about particular projects or families of projects. The register was launched in December 2012.

14. See Annex B for the full explanation of the arrangements.

## IV. Future directions

15. As with any government wide initiative, there is much work that still needs to be done so that the public can be assured that integration of Australian Government data is conducted safely and securely.

### A. Education

16. Once the final framework is in place, one of the main streams of work will be to educate government staff about the framework and to ensure adherence to the framework. If some departments and their staff are not aware of the framework, how it works, and the expectations it outlines, there is a risk that they will integrate data outside of the framework. This will mean that the aim of the framework, to ensure data is integrated securely in a way that respects the trust of the public, will not be realised. Therefore, as part of the agreed governance arrangements, education and training has also been outlined as a key part of the framework and will need to be progressed in 2013 and 2014.

17. There has been and will continue to be much work conducted by the ABS to assist in enhancing agencies' capabilities in undertaking data integration activities. Many different methods can be used to link data depending on the quality of the data and differences between the data sets being integrated. Therefore, assisting agencies in adopting best practice techniques for various data linking projects will play a key role in improving the quality of the results found through research using linked data. For example, even in cases where direct identifiers are present, such as names and addresses, inaccuracies or inconsistencies in the way the data is reported can make integration activities difficult. The ABS has made substantial investment over a number of years to develop capability and systems that accommodate a range of data linking methods for use in different situations, including the development of databases to hold common name and address variants to increase the likelihood of correct record linking across datasets holding such identifiers.

### B. Dissemination

18. Ensuring the confidentiality of integrated data from a dissemination perspective may pose new challenges. Standard practices, such as the release of confidentialised unit record

data, may no longer be practical, particularly where multiple data sources have been utilised in an integration activity. Further work is needed to protect confidentiality whilst providing access to micro-data. In the ABS, large investments have been made over a number of years to develop dynamic analysis and tabulation tools. These tools enable users to undertake custom queries/analysis with in-built confidentiality routines to ensure analytical outputs produced will maintain the confidentiality of respondents' information. The use of such tools will also need to be considered by other agencies in the future.

### **C. Expansion of the Commonwealth arrangements**

19. The current framework only covers Commonwealth agencies. However, state and territory governments also hold large amounts of data that could be unlocked for data integration projects. They also have similar responsibilities to maintain the privacy and confidentiality of personal and organisational information and will suffer the consequences of public distrust if a data breach occurs. Any state, territory, or local government breach is also a risk for the Commonwealth government as many Commonwealth agencies rely on the compilation of data collected by other tiers of government for monitoring, policy development and research. Therefore, in the future, the framework will be extended to allow state and territory government agencies to become accredited Integrating Authorities. This will allow 'high risk' projects using a mix of state/territory and Commonwealth data to be conducted by organisations that have met the Commonwealth accredited IA standards. However, it is envisaged that accreditation of state and territory organisations will not commence until the interim arrangements regarding the accreditation process have been reviewed and finalised.

### **D. Data quality**

20. Underpinning all data integration, and the value it can bring to assisting in answering complex research and policy issues, is the quality of the datasets being integrated. Therefore, future work for all IAs, and particularly the ABS, will include continuing to encourage the improvement of the quality of the data collected through government administration processes. This includes encouraging consistency across the country in areas such as collection methodology. For data collected to date, there will also need to be the development of quality frameworks to understand and articulate to users the quality of the data including concepts, coverage and accuracy.

### **V. Conclusion**

21. Although there is still work to be done to ensure that the Commonwealth Data Linkage Framework is finalised and implemented, the work conducted to date has provided clear guidance to all Commonwealth government agencies about the expectations placed on them when integrating data. For example, the work invested in the development of the framework indicates to agencies that there is a Commonwealth wide expectation to utilise the data they collect as a strategic resource to assist in solving complex and 'wicked' problems. The framework also ensures that all agencies can respond to the need for data integration projects in a consistent way, as well as providing them with a common language and the confidence to discuss the possibilities of sharing their data with other agencies (where relevant legislation allows) for safe and secure data integration.

22. It is hoped that these benefits, along with the clear responsibilities the framework sets out for Integrating Authorities in protecting privacy and confidentiality, will ensure the

development of a vibrant array of rich datasets from which academics, researchers and the Australian Government can draw new insights and future policy directions.

## **Annex A**

### **High level principles for statistical data integration across Australian Government: endorsed by portfolio secretaries 3 February 2010**

#### **A. Principle 1 – responsible agencies should treat data as a strategic resource and design and manage administrative data to support their wider statistical and research use**

1. Administrative data represents a public asset that requires protection and management for appropriate purposes. When designing and managing administrative datasets, the responsible agency should consider the potential statistical value of the datasets for public good, both in terms of use by their own agency, and use more broadly.
2. Administrative data cannot be used for statistical purposes if this contravenes legislation or any commitment made to data providers, or the data is commercial in confidence. Nor should it be used for statistical purposes if this use clearly threatens the integrity of the administrative data.
3. Where administrative data is likely to have high public value for statistical use, those providing data should be informed of the potential for statistical use at the time of data collection. Where historical data has been collected without providing this information it should still be considered for statistical use, but not if this is prohibited by commitments made to providers at the time of collection.
4. Where administrative data is likely to have value for statistical use, efforts should be made to maximise that value through good data management. This includes the use of standard definitions and classifications and the maintenance of appropriate metadata, including quality attributes of the data.
5. Where data is sought for statistical purposes, consideration should be given to using existing administrative sources in preference to imposing additional load on providers through the institution of a new statistical collection.
6. The statistical and research value of administrative data should be maximised, within legal and practical constraints, by granting broad access for research purposes, to data that is not likely to enable identification. Commonwealth administrative data should not generally be withheld from research for reasons of Intellectual Property.

#### **B. Principle 2 – agencies responsible for source data used in statistical data integration remain individually accountable for their security and confidentiality**

7. Each responsible agency for source data:
  - (a) Must agree on mechanisms to achieve adequate control and manage risks appropriate to their own situation. For some these mechanisms may include the use of particular institutional arrangements with trusted institutions, the use of specified standards and audits against those standards, and the potential application of sanctions;



(b) Will need to agree to the nature of valid uses that can be made of the integrated datasets and the approval mechanism to be applied to applications to use the datasets, as well as any control mechanisms to be applied to such use;

(c) Will need to manage the potential increase in identifiability of data for which they are responsible when it is used in conjunction with data from other sources. It will need to agree on mechanisms by which it can assure itself that outputs from the statistical data integration are not likely to enable the identification of individuals or businesses;

(d) Will need to agree on the final content of any new data integration proposal, or any material change to an existing data integration proposal as part of the approval process. They must be kept informed of, and agree, more minor proposed changes to existing proposals.

8. Where an agency does not agree to the use of its source data in a statistical integration proposal, that data will not be included. For example this might occur if the proposal threatens the integrity of the administrative data.

### **C. Principle 3 – a responsible ‘integrating authority’ will be nominated for each statistical data integration proposal**

9. An integrating authority must be identified for each statistical data integration proposal. This authority will be held responsible for the sound conduct of the statistical data integration proposed, in line with the agreed requirements of the responsible agencies.

10. Although the integrating authority is the single organisation ultimately accountable for the Statistical Data Integration project, it may work with a network of agencies to achieve the data integration, for example it might use another agency to undertake linkage or to support dissemination.

11. The integrating authority will ensure appropriate governance is in place including:

- (a) An open approval process is followed;
- (b) Documentation of the proposal;
- (c) The impact on privacy;
- (d) Risks have been assessed, managed and mitigated;
- (e) The expected costs and benefits; and
- (f) The outputs.

12. A family of data integration projects using the same source datasets, for similar purposes, with the same integrating authority, may be treated as a single program for the purposes of the approval process.

13. The integrating authority will be responsible for the ongoing management of the integrated data, ensuring it is kept secure, confidential and fit for the purposes for which it was approved.

14. If it is an ongoing project, the integrating authority will be responsible for initiating and managing its regular review, in consultation with source data agencies.

**D. Principle 4 – statistical integration should only occur where it provides significant overall benefit to the public**

15. There should be a demonstrated ability to produce significant outputs from the integrated dataset.

16. There should be an independent assessment of the balance of public good against the privacy imposition and risks to confidentiality. Examples include community representation on the steering committee, the use of an ethics committee, or the use of an advisory committee with community representation and the ability to report independently of the agencies involved in the proposal.

17. Ongoing programs should be reviewed on a three yearly basis to ensure a continuing overall benefit.

**E. Principle 5 – statistical data integration must be used for statistical and research purposes only**

18. Statistical data integration must not be used for non-statistical purposes requiring the identification of an individual person, household, family or business, for example the delivery of services to particular individuals, individual compliance monitoring, client management, incident investigation, or for regulatory purposes. However the insights gained through statistical and research outputs are expected to improve processes in these areas.

19. There must be no feedback of information relating to individuals or individual businesses, from the statistical data integration project back to the originating administrative sources, unless that feedback was derived from a single source and is returning the same data to that source.

**F. Principle 6 – policies and procedures used in data integration must minimise any potential impact on privacy and confidentiality**

20. Operational, administrative and personal identifiers should be removed from datasets as soon as they are no longer required to meet the approved purposes of the statistical data integration.

21. Where identifiers need to be retained, for example for longitudinal studies, they should be kept separate from the integrated dataset.

22. The number of unit records and data variables to be included in an integrated dataset should be no more than required to support the approved purposes.

23. The type of matching used (exact, probabilistic or statistical) should be chosen as the minimum needed to support the approved purposes, and the range of attributes used to establish a common identity should be the minimum necessary for the linking operation to succeed.

24. Access to potentially identifiable data for statistical and research purposes, outside secure and trusted institutional environments should only occur where:

- (a) Legislation allows;
- (b) It is necessary to achieve the approved purposes; and
- (c) It meets agreements with source data agencies.

25. Risks of indirect as well as direct identification should be carefully managed when data is disseminated outside secure and trusted institutions, particularly in terms of units with unusual characteristics. This management must take account of the potential increase in identifiability of one set of data when combined with another set. It might involve strict data use licensing conditions, reducing detail, perturbing data, or seeking the consent of the individual or business involved to release potentially identifiable data, the last of these being most likely in the case of business data.

26. Once the approved purpose of the project is met, the related datasets should be destroyed, or if retained, the reasons for and necessity of retention documented, and a review process set up. If such retention was not part of the initial approval process, re-approval of the decision to retain is required.

27. Archiving of statistically integrated data sets should be restricted to confidentialised datasets.

#### **G. Principle 7 – statistical data integration will be conducted in an open and accountable way**

28. The main elements are:

(a) Ensuring the governance of the project is managed in an open accountable way;

(b) Ensuring stakeholders and the community are kept informed of any statistical data integration project undertaken, by publishing appropriate details of the project such as: the datasets, the purpose, provision for access, use made of the dataset, the makeup and role of any advisory body or steering group, the role of involved institutions, the approval process, and the review process;

(c) Ensuring that an appropriate privacy impact assessment has been conducted;

(d) Ensuring that the agencies responsible for source data and the data integration, agree on audit schedules.

## Annex B

### **Data integration involving Commonwealth data for statistical and research purposes: governance and institutional arrangements, endorsed by portfolio secretaries 6 October 2010**

#### **A. Cross portfolio data integration oversight board**

1. A high level **Cross Portfolio Data Integration Oversight Board (the Board)** will be established by early 2011 to oversee the development of a cross government environment for data integration involving Commonwealth data for statistical and research purposes that is safe and effective. The Board will be chaired by the Australian Statistician and membership will initially include the heads of the Department of Health and Ageing; the Department of Families, Housing, Community Services and Indigenous Affairs; and the Department of Human Services.

2. The **role** of the Board will be to:

(a) Provide strategic and collaborative leadership, support effective governance and help manage the risks of particular data integration projects;

(b) Help manage the systemic risk associated with conducting multiple data integration projects involving Commonwealth data through assessment of proposed risk mitigation strategies, and the provision of advice; and

(c) Endorse any changes or additions to the overall environment, including amendments to the principles or guidelines, or the development of new general tools to support integration or safe access to integrated data for statistical and research purposes.

3. **Activities** of the Board will include:

(a) Advising on data integration projects assessed as being of **high risk** and with the potential to significantly impact information related activities across government. Where the Board advises on amendments to, or discontinuation of, a particular project, the relevant custodians and Integrating Authority will need to consider these views in deciding whether to proceed with the proposed project;

(b) Working with agencies to help ensure the systemic risk associated with high risk projects is adequately managed. It will do this by reviewing projects involving data integration of Commonwealth data for statistical and research purposes deemed by agencies to pose a significant level of systemic risk, and advising on the sufficiency of the intended risk mitigation strategies;

(c) The review of any adverse incidents of high public concern relating to data integration involving Commonwealth data for statistical and research purposes, and seen as having a likely systemic impact on public trust in government use of data;

(d) Consultation with the Office of the Privacy Commissioner and the Office of the Australian Information Commissioner, as required. This allows for the Commissioners to be at arms-length from any decisions made by the Board. The Board is also free to consult representatives from the research community or other parties as required.

## **B. Accreditation of authorised integrating authorities for high systemic risk projects**

4. An essential pillar of establishing a safe and effective environment for statistical data integration is the nomination of an agency as the authorised Integrating Authority for each statistical data integration proposal. The Integrating Authority will be responsible for the sound conduct of the integration project and may work with other agencies to achieve components of the statistical data integration proposal for example it might use another agency to undertake linkage or to support dissemination. The integrating authority has overall responsibility to ensure that risks have been assessed, managed and mitigated throughout the duration of the project, including regular reviews of ongoing projects.

5. As described by Principle 3 of the High Level Principles, the Integrating Authority will ensure appropriate governance is in place for the statistical data integration proposal including: using an open approval process; documenting the proposal; considering the privacy impacts, examining the expected costs and benefits of the proposal and considering the access arrangements and dissemination plans. The Integrating Authority will be responsible for the ongoing management of the integrated data, ensuring it is kept secure, confidential and fit for the purposes for which it was approved.

6. Integration authorities are responsible for the implementation of the data integration project, and the management of the integrated datasets throughout their life cycle, ensuring full compliance with the commitments made as part of the project approval.

7. A key requirement of Integrating Authorities is that, to the extent that the data they deal with involves identifiable information, they be in a position to comply with the requirements of the Privacy Act 1988 (insofar as information about individuals is involved) and secrecy provisions generally (insofar as information with respect to the affairs of any third party, corporate or individual, is involved). This may require either the consent of the individual to the particular use or disclosure for Privacy Act 1988 purposes, or an overriding public interest test certified in accordance with the relevant secrecy provision.

8. Integrating Authorities will only be established at the initiative of an interested agency. Such agencies may wish to involve their respective Ministers in the course of preparing a proposal for accreditation to the Board. An Integrating Authority could also be established administratively within a Department or other agency, that is, it would be part of an agency subject to the provisions of the Privacy Act 1988.

## **C. Accreditation process**

9. For data integration proposals considered by custodians to pose a high systemic risk, nomination of an authorised and accredited Integrating Authority is required. An **accreditation process** will be established through the Board to enable the endorsement of authorised and accredited Integrating Authorities with the capacity to deal with high risk data integration projects or families of projects involving Commonwealth data.

10. Data integration projects involving Commonwealth data for statistical and research purposes judged to pose a high systemic risk will need to be undertaken with particular care to help mitigate this risk. This will require a high level of relevant expertise, a strong understanding of, and capability for, maintaining security, as well as a consistently high standard of behaviour by all employees based on a strong culture, and set of values. To ensure effective use of specialist skills and infrastructure, the number of accredited integrating authorities is expected to be relatively limited.

11. An **accreditation process** including interim arrangements will be proposed through cross government consultation. It is expected that interim arrangements will be proposed for discussion and endorsement by the Board in early 2011 with final arrangements agreed early 2012.

**D. Guidelines for all data integration involving Commonwealth data for statistical and research purposes**

12. A comprehensive set of **Guidelines** describing best practice for data integration projects involving Commonwealth data will be jointly developed by a cross government working group and approved by the Board. These Guidelines will build upon already available guidelines (e.g. *National Statement on Ethical Conduct in Human Research*, related guidelines of the Office of the Privacy Commissioner, and agency specific guidelines and proformas).

13. The Guidelines will be considered in conjunction with legislation relevant to the data custodians, data providers and the Integrating Authority.

14. The Guidelines will cover such areas as:

(a) Privacy including Privacy Impact Assessments and the protection of privacy in medical research<sup>5</sup>;

(b) Approval;

(c) Registration;

(d) Agreements - between data custodians and the Integrating Authority; or with end users;

(e) Responsibilities of Custodians and Integrating Authorities including accreditation criteria and accreditation process;

(f) The handling of identifiers and application of the separation principle, for example, the separation of identifiers and key demographic data used for linking (e.g. date of birth) from the individual content or event information (e.g. clinical or benefit information);

(g) The minimum standards for the secure management of data during and after integration, including the exchange of sensitive data across agencies;

(h) The minimum standards for consent to access Commonwealth government data;

(i) Confidentialising of integrated data sets and research outputs;

(j) Managing access and use of integrated data sets;

(k) The minimum standards for data destruction or review.

15. The Guidelines will be developed progressively, commencing in 2011 with guidelines relating to Integrating Authorities. They will be made publicly available on an internet site, and maintained by the Secretariat to the Board.

---

<sup>5</sup> *Privacy Act 1988* Section 95 Medical Research Guidelines available at [http://www.austlii.edu.au/au/legis/cth/consol\\_act/pa1988108/s95.html](http://www.austlii.edu.au/au/legis/cth/consol_act/pa1988108/s95.html)

## **E. Education and training to support data integration projects using Commonwealth data**

16. An **education and training strategy** will be developed and implemented to support the Guidelines for data integration projects involving Commonwealth data for statistical and research purposes.

17. The Guidelines and the education and training strategy will be targeted to key stakeholders including government agencies (acting variously as custodians, Integrating Authorities, researchers, Privacy commissions), the private sector (as custodians or researchers), the academic/research community and the public.

18. The purpose of the education and training strategy will be to build an understanding of:

(a) The underpinning values of a safe and effective environment for data integration projects;

(b) Key issues in managing privacy and confidentiality in data integration project;

(c) Key legislative requirements within the management of the data integration project including maintaining privacy and confidentiality;

(d) Sanctions that apply for non-compliance with legislative requirements or other institutional requirements;

(e) Support available in terms of specialist advice, knowledge and tools.

19. Following its development and finalisation by October 2013, the education and training strategy will be publicly available on an internet site and maintained by the Secretariat to the Board.

## **F. Network of statistical liaison officers in agencies**

20. Statistical Liaison Officers (SLOs) have been established in many Commonwealth agencies to provide a central contact of information and expertise in statistical collections involving businesses. Statistical Liaison Officers are part of an Australian Government Initiative established by the Prime Minister in 1997 to reduce duplication, minimise the burden on business, and ensure that statistical data collections involving businesses are fit-for-purpose.

21. The role of the existing network of **statistical liaison officers (SLOs)** across Commonwealth agencies will be expanded to provide a central contact of information and expertise in data integration projects involving Commonwealth data. This network will be able to share information on good practice and help minimise duplication.

## **G. Register of data integration projects involving Commonwealth data conducted for statistical and research purposes**

22. Consistent with a more open government engaging online with the public, a web-based **register**, will be established to include a description of data integration projects involving Commonwealth data for statistical and research purposes and a **public feedback mechanism** to enable members of the public to register support or concerns and make suggestions about particular projects or families of projects.

23. The registered entry will describe the project purpose and data sources, recognising that specific details on individual use will generally not be made publicly available for in-confidence government business. For similar reasons, actual data will not be included in the registration entry. A format for such submission will be developed across government and made available through self-help mechanisms (email, SLO network, internet site).
24. The Integrating Authority nominated for each project will be responsible for submitting an entry to the register once the project is finalised.
25. The register will be available through an appropriate whole-of-government web site, along with related information, including the High Level Principles, the Guidelines and access to the public feedback mechanism. It is expected that development of the register will commence in 2011 with trialling through 2012 and final release by 2013.
26. The Secretariat will maintain the register and associated internet site.

## **H. Secretariat to support cross government activity and infrastructure**

27. A small **Secretariat** will be established by early 2011 and based in the Australian Bureau of Statistics, to support the Board and its ongoing activities.
28. The Secretariat will support the Board by:
  - (a) Providing administrative support as needed to facilitate meetings, maintain centralised documentation and undertake investigation or development work as directed by the Board;
  - (b) Assisting the cross government development of the Guidelines, the development and, in part, the delivery of the education and training strategy, and expansion of the network of Statistical Liaison Officers;
  - (c) Developing and maintaining a web-based public register of the description of projects and the provision of a public feedback mechanism, in conjunction with other agencies;
  - (d) Monitoring the risk assessment of projects, including legal and systemic risks, advise project managers where further advice has been requested, and identify instances where apparently high risk projects have not been assessed as such, initiating discussion with the agency involved in the first instance to resolve concerns, and as required, escalating the assessed level of risk, for Board consideration;
  - (e) Providing a contact point in government for issues relating to data integration involving Commonwealth data for statistical and research purposes, whether coming from data custodians, Integrating Authorities, researchers or the public;
  - (f) Following development of the public register, releasing an annual report on key aspects of data integration activity involving Commonwealth data, including a summarised report on public feedback.

### **1. Transition arrangements**

29. Transition arrangements will apply during the early stages of the implementation of the high level principles and the governance and institutional arrangements outlined above. Transition arrangements include the following:
  - (a) The Board and Secretariat will be established in early 2011 to commission cross government work to develop and begin promulgating the guidelines and the education and training strategy, the use of statistical liaison officers, as well as accreditation criteria for Integrating Authorities;



(b) In early 2011, it is expected that a small number of agencies will be accredited as authorised Integrating Authorities particularly for high risk projects. This will be an interim accreditation, subject to review and finalisation of the formal accreditation process;

(c) The education and training strategy will be developed to provide a range of self-help tools to inform stakeholders. This will be expanded over the following two years to include seminars, workshops, and conferences. The strategy will also be designed to develop the skills of Integrating Authorities;

(d) New data integration projects which commence prior to the development of the Guidelines are expected to be set up to conform to the high level principles;

(e) High risk projects should be notified to the Secretariat for review by the Board. The project should be fully developed before consideration by the Board. However, in the initial stages, it may be useful for those developing the proposal to liaise with the Secretariat during its development. Where a new, potentially high risk project has been substantially developed, and committed timeframes do not permit review by the Board, the Board should be notified of the project and it should be otherwise treated as a project currently underway;

(f) Those projects, including families of projects, which are underway as at October 2010, and which are expected to be completed by October 2012, will not be subject to the new arrangements, although project managers are encouraged to consider the conformity of the project to the principles and to seek advice from the Secretariat if there are aspects of the projects that may be of concern;

(g) Those projects, including families of projects, which are underway, and are expected to continue beyond October 2012, will need to become aligned with the principles, and with the proposed governance and institutional arrangements. It is proposed that such projects are brought into alignment by October 2012.

---