



A GOOD PRACTICE GUIDE TO SHARING YOUR DATA WITH OTHERS

Version 1, November 2009

Government organisations are increasingly interested in sharing data, between and across all levels of government, to realise the full potential of data, and build a comprehensive and coherent statistical picture of the economy, society and the environment. Data sharing allows maximum use of data for statistical purposes, thus enhancing the decision-making capability of governments and communities. It is an important ingredient for supporting evidence-based policy and decision-making.

This guide provides a non-technical introduction to data sharing and describes the key concepts and components of data sharing arrangements and agreements. It is designed to assist:

- government data custodians, data providers and data users to think and communicate about data sharing between and across all levels of government
- government agencies understand data sharing benefits, barriers, principles and enablers
- those involved in establishing or negotiating data sharing arrangements achieve successful outcomes
- agencies to work together to achieve data interoperability through the exchange of data and information
- data custodians, data providers, data users and communities achieve transparency when sharing data
- staff working on, or contributing to, data sharing projects by promoting an understanding of data sharing, the key drivers, benefits and barriers, typical forms, and some mechanisms for supporting data sharing between agencies

The guide provides a basic framework for data sharing agreements, including key factors to consider when developing data sharing agreements and how these relate to the agency, staff and data. The guide draws on the experiences of those undertaking data sharing for statistical purposes and from nationally endorsed frameworks and strategies such as the [Information Interoperability Framework: Sharing Information across boundaries](#) and the [National Government Information Sharing Strategy](#).

Note: This guide contains general advice and guidance on data sharing; however, it is not an exhaustive description of data sharing arrangements and issues. When considering data sharing, we recommend that you seek appropriate professional advice from areas including the in-agency policy/legislation section, the Government Solicitor, and the Privacy Commissioner.

BACKGROUND

A key theme of Australian Government policy is that agencies should work together to better respond to complex policy challenges and to improve the delivery of services to Australian citizens. Agencies are increasingly required to reach across portfolio boundaries to find collaborative, networked and multi-channel approaches to delivering information and services. This expectation is clearly articulated in the Fourth Management Advisory Committee (MAC4) 2004 report, [Connecting Government: Whole of Government Responses to Australia's Priority Challenges](#).¹

The Council of Australian Governments (COAG), through its initiatives, has affirmed its support of data sharing arrangements between jurisdictions to enable a better understanding of issues and enhance policy development and evidence-based decision making. As part of this, performance indicator data have been linked to a number of National Agreements in key policy areas signed by COAG. COAG recognises that through cooperative action, including data sharing, improved decision-making and implementation of services and programmes can be achieved, and data gaps identified and addressed.

The [National Government Information Sharing Strategy](#) (NGISS), endorsed by the COAG Online and Communications Council (OCC) in December 2008, provides the direction for all levels of government to adopt the nine information sharing principles outlined in the strategy. The NGISS principles include promoting information re-use and acting collaboratively. A detailed NGISS implementation plan is being developed in 2009.

There is a breadth of untapped information that is generated and held across all levels of government - this information is a rich statistical data source. At the same time, many agencies recognise the costs associated with data collection and use. By reusing existing information for statistical purposes, increased value and use of the information can be made to support government policy/program development, implementation and review² and respondent burden can be minimised. Data interoperability and comparability can be progressed by sharing and re-using existing data held across all levels of government.

WHAT IS DATA SHARING?

Data sharing is the transfer of data between two or more parties. It has been taking place for many years across governments, research bodies, business and other bodies, assisting informed decision-making, research and discussion within governments and the wider community.

Data sharing can take many forms, from sharing metadata (information about data), to sharing aggregate data, to sharing survey or unit record data or a complete administrative dataset. Sharing metadata exposes the availability of data and is a key first step in sharing any type of data. Indeed, sharing metadata may be all that is possible in some cases. In other cases, sharing aggregate statistical data is all that is required by government to inform decision-making.

Data sharing can be one-way, two-way, or involve multiple parties or distribution into the public domain. In many cases, information requested for data sharing purposes is not readily available in the public domain.

Data sharing is not data linking

Data linking is also known as data matching, record matching, duplicate detection and deduplication. For the purposes of this guide, a distinction is made between data linking and data sharing. The key difference is data linking brings data together from across datasets in a way that people or events are brought together (linked). More specifically, data linking is where records from two or more datasets are brought together in a way that joins separate data records belonging to the same person or business.

Data linking is an activity that may be applied to shared data where the level of detail and legal conditions allow it. Data linking takes data sharing one step further. The legal and privacy issues for data linking go beyond those related to data sharing. In this guide, data sharing does not include data linking. Further information on data linking is available from the Australian Bureau of Statistics.

WHAT CAN DATA SHARING LOOK LIKE?

Some examples of data sharing include:

- joint publications, for example, *The Health and Welfare of Australia's Aboriginal and Torres Strait Islander Peoples* (ABS cat. no. 4704.0) is a joint publication of the Australian Bureau of Statistics (ABS) and the Australian Institute of Health and Welfare (AIHW).
- sharing metadata to enable data to be coded to another agency's classifications or standards, for example, some ABS classifications and standards are made available through AIHW's Metadata Online Registry, [METeOR](#) for use by users of this repository.
- sharing metadata to increase the visibility, or expose the existence, of data, for example, the Transport Metadata Portal, a pilot metadata website developed to support the National Transport Policy Framework.
- data pooling, for instance combining data from multiple collections, for example, the *Labour Force Characteristics of Aboriginal and Torres Strait Islander Australians, Estimates from the Labour Forces Survey* (ABS cat. no. 6287.0) which uses data from several collections throughout the year to compile the estimate.
- modelling, for example, Commonwealth Treasury's *Macroeconomic (TRYM) Model* uses ABS information to model the Australian economy and macroeconomy. For more information see www.treasury.gov.au.
- quality assessment or validation of another data set, for example ABS collections confront their data with other data held by the ABS and other organisations to ensure coherency. For more information see *Information Paper: Quality Dimensions of the Australian National Accounts* (ABS cat. no. 5216.0.55.002).
- analysis and value-adding, for example, *CData Online* (ABS cat. no. 2064.0) was jointly developed by the ABS and Space-Time Research. This online tool combines 2006 Census data with web graphing and mapping capability, thus enhancing the representation and visualisation of the data.
- bringing together statistical information from a range of data sources, for example community indicator frameworks such as [Regional Spotlights](#), to provide a practical framework to understand joint issues.

WHAT ARE THE BENEFITS OF SHARING DATA?

Data sharing is one means of facilitating the provision of trusted statistics to answer the questions that are important to Australians, their families and their communities, resulting in a better life through better public and private decisions. It can assist better policy and business outcomes for both producers and users of data.

Data sharing can enable a more complete statistical picture to be drawn around specific areas of interest due to the availability of a greater breadth of data for analysis. Sharing data held by government also enables Australia to realise the potential of its existing data. It reduces costs of data collection and management by reducing duplication and effort. It also reduces the burden placed on individuals, households and businesses to provide information to government agencies.

Data sharing can make data sources and datasets more useful by improving metadata management and sharing. The application of common metadata across datasets can increase comparability and consistency, helps data providers to support comparable datasets and helps data users to access similar information across multiple data sources.

In summary, data sharing makes sense because it can:

- provide a more cost-effective solution when finding, using or exposing data
- improve the usefulness of available data sources and datasets to data producers and users
- improve decision-making, policy and business outcomes for Australians due to the increased availability of information
- improve relationships between data custodians, providers and users through the identification of common interests and access to information
- maximise the awareness and use of existing data
- reduce respondent burden

WHAT NEEDS TO BE DONE TO SHARE DATA?

Data sharing across government requires collaboration and commitment to data sharing by agencies. To share data, agencies must also agree on standards for managing and sharing data and establish a clear purpose for sharing data. Data sharing arrangements on any scale and frequency are best supported by the establishment of a data sharing agreement between the agencies.

DATA SHARING AGREEMENTS

Data sharing arrangements should be supported by a data sharing agreement. Data sharing agreements clearly state the terms and conditions of use for the data. There are many types of data sharing agreements, including memorandums of understanding, contracts, deeds, letters, undertakings, licences, head of agency agreements, and so on. It is important that the right type of agreement is used for the right circumstance, for instance:

- a letter of exchange might be used for a one-off sharing arrangement
- a memorandum of understanding might be used for on-going data sharing between agencies

- a contract might be needed if money is exchanged

The agreement documents the relationship between parties, which data components and elements are to be shared, and how the data may be used. Following are some key components and elements to consider in a data sharing agreement.

Data Sharing Agreements - the key components and elements		
A	Aims and purpose	Having a clear and common understanding of the aims and purpose of a data sharing arrangement is a key component of a data sharing agreement.
B	Data definition	Defining the data being shared, including metadata and formats, and the terms of the data sharing agreement are important elements of a data sharing agreement.
C	Legal restrictions	Ensuring data sharing arrangements comply with legislation and other consent based restrictions are a key element of data sharing. The legislative and consent based restrictions should be documented in a data sharing agreement.
D	Governance	The governance arrangements are an important element of all data sharing agreements and cover all aspects of the data sharing including management and termination.
E	Access issues	Access issues are a key component of data sharing agreements because they detail how the data will be accessed and shared. Detailing the issues and processes for sharing the data is an important element of ensuring compliance with legal and legislative issues and maintaining community trust.
F	Data quality	Provision of information about data quality including, the data's "fitness for purpose" is an important element of data sharing agreements because it facilitates the understanding of the data by the agencies and enables transparency.
G	Data management	Data management encompasses the systems and processes that ensure data integrity and security, including metadata, data security and access registers. This is an important component of a data sharing agreement.
H	Costs	In some cases, costs may be associated with a data sharing arrangement. The details of any costs, including a payment schedule are a key component of a data sharing agreement.

A. Aims and purpose

Develop a common understanding about the purpose and expected benefits of the data and the data sharing arrangement. In establishing an understanding of the aims and purpose consider:

- Why is the data being shared? In what context is it being shared?
- What will the data be used for?
- How will the data be used?

Why? Communicating and documenting the aims for all parties promotes understanding and commitment to the data sharing arrangement. Clearly identifying the aims makes it possible to periodically review the practical outcomes of the sharing arrangement to ensure the aims are being achieved, and to refine the arrangement if they are not. It also enables assessments of any proposed new uses of the data, or changes to how the data is provided, to ensure the data is consistent with the original intent and spirit of the data sharing arrangement.

B. Data definition

Define the data that is being shared including the processes that create the data (e.g. commercial, administrative, statistical), the frequency of collection and provision, and the associated metadata.

Consider:

- How was the data created or generated?
- For what purpose was the data collected, created or generated?
- Is the data final or subject to revision?
- What standards and classifications are used in data collection, creation or generation?
- What descriptive or definitional metadata is available for the data?
- How often is the data collected or generated?
- In what format will the data be shared?
- What variables will be shared?
- How often will the data be shared (e.g. daily, monthly, quarterly, annually)?

Define the end use of the data once it is shared. For example, establish whether the data will be used in a commercial context, published, or used with other data.

Consider:

- How will the data be used by the data requestor?
- For what purpose has the data been requested?
- How will the data be interpreted and/or analysed?
- What is the planned end-use of the data?
- Will the data be combined or linked with any other data?

Why? These definitions make it clear what data is being shared and how it will be used. This promotes understanding and adds value to the data sharing arrangement as all parties can understand how and when it was created and structured and the intended use of the data by others. It also enables data users to establish interaction and comparability with other datasets.

C. Legal restrictions

Take the time to ensure the data sharing arrangement complies with legislation (including copyright and privacy laws, and agency-specific legislation and policy), and give due consideration to other legal issues and consent based restrictions. It is important to note, laws may differ across jurisdictions. As agencies can operate under different confidentiality and secrecy provisions and legislation, appropriate consideration should be given to ensuring that the data sharing arrangement and agreement complies with the agency-specific legislation and provisions. Also, in some cases, there may be a need to obtain ethical approval for any piece of

work that implicates the involvement of others or has the capacity to interfere with or make a difference to their lives prior to data sharing.

Often data compilations are protected by copyright, if so; the permission of the copyright owner is required. Creative Commons licencing is increasingly being adopted by government agencies to manage licencing. Creative Commons offers flexible licence management and encourages collaboration and innovation in the further use of Creative Commons licensed information, while still offering legal protection to the original author. For more information on Creative Commons see www.creativecommons.org.au and the Government Information Licensing Framework see www.gilf.gov.au.

Consider:

- Who owns the copyright?
- Are the copyright terms "All Rights Reserved" or "Some Rights Reserved"?
- Are any licenses required to access the data?
- What policy or legislation applies to the data?
- Are there privacy concerns around the data?
- Is the data identified or identifiable?
- Who will own the Intellectual Property?
- Who will have custodianship of any new datasets created?
- What legislation will the datasets be held or protected under?

Why? Honouring and enforcing legislative and consent based restrictions, including copyright, are legal requirements. Further, they are important in maintaining community trust and provide clear guidelines on what can and can't be done by parties. Failure to comply could result in negative consequences for all parties, other data collections, and the government.

D. Governance

Establish governance arrangements for all aspects of the data sharing, including the process and operational management, and maintaining and terminating the agreement. As part of developing a governance structure, it is also important to identify any risks associated with the data sharing arrangement and incorporate a risk mitigation strategy. In developing governance arrangements consider:

- Who owns the data? Who is the data custodian?
- What legislation applies?
- How does the agreement meet the obligations of the relevant legislation?
- Who are appropriate signatories for the agreement?
- What are the roles and responsibilities of all parties?

- What feedback mechanisms will be established?
- How will the agreement be maintained?
- What review and audit processes will be put in place?
- What is the process for making variations to the data sharing arrangement or agreement?
- How may the agreement be terminated?
- What are the expected and agreed service levels around data delivery?
- What are the risks associated with the data sharing arrangement and how can these be mitigated?

Why? The arrangement and agreement have a lifecycle, from creation to actual use and termination. Having clear governance arrangements allows the data sharing arrangement and associated relationships to be more easily managed and maintained. It also establishes roles and responsibilities of all parties involved including legislative or legal obligations. Having a well-governed data sharing arrangement encourages real interaction between the parties based on common understanding and mutual obligation.

Ensuring the agreement is signed by the most appropriate person on behalf of each party gives it authority. To determine the appropriate signatories for an agreement, consider the costs, legal or legislative issues, the complexity and significance of the agreement, and the relationship between parties.

E. Access issues

Consider access issues to the data and make it visible in the agreement, including detailing the issues and processes for this arrangement. For instance, data users may be required to sign an undertaking or license before being able to use the data.

Consider:

- Who will have access to the data?
- Is a license required to access the data?
- Is the data identified (or identifiable)?
- How long is access to the data required?
- Do individual users of the data require any type of approval from the data owner to access the data?
- How will compliance with the access requirements be audited?
- When will the access lists be reviewed?
- Who will review the access lists?
- Does any legislation or policy impact on the provision and use of the data?

Why? Access to data often depends on legal and legislative issues, for example privacy, confidentiality, license arrangements and legislation, such as the *Census and Statistics Act 1905*. In some cases, data may only be shared with specified individuals, in other cases, access can be granted to an entire organisation. These access requirements should be observed by all parties, with all involved understanding why the access requirements exist, the implications of non-compliance and the signing of appropriate documents (e.g. undertakings or licences).

Also consider how the data will be accessed and the processes for obtaining and maintaining the access and make it visible in the agreement. This may include the technological platforms to ensure the data is protected and data is shared with those who are authorised to use it.

Consider:

- How will the data be transferred from one agency to the other?
- What software is required for access to the data?
- Is the data classified? i.e. at what level is the data classified (unclassified, in-confidence, protected, etc)?
- Who is authorised to see and use this data? Do staff require any special clearance?
- How will the data be stored or secured by the data requestor?
- Is any education or special support required for use of the data?
- Will any restrictions be placed on access to the data?
- How will the data be disposed of?

Why? Access to data may be dependant on having access to a specific system or software, or the ability to transfer data via a specific method. Data may only be able to be shared via a specified technological platform or software. Also, data may need to be transferred between parties via specific channels, including designated encryption and transfer methods. All technological aspects for data sharing should be included in the agreement, including any implications for non-compliance.

It is important to note that activities which breach an agreement may result in consequences or penalties for the individual or organisation in breach of the agreement and may jeopardise any further data sharing arrangements and community trust.

F. Data quality

Provide information about data quality so it is clear how authoritative, final and current the data is. Ideally, this information should be presented in a Data Quality Statement and be provided with the data. The Data Quality Statement should include information on institutional environment, relevance, timeliness, accuracy, coherence, interpretability and accessibility. For more information on data quality, refer to the [ABS Data Quality Framework](#) (ABS cat. no. 1520.0).

Consider:

- Is a data quality statement available?
- Is there enough information available to assess the data's fitness for purpose?

- Are technical and explanatory notes available for the data?
- What methodology was used to create or generate the data?
- What processes are used to monitor and assess data quality?

Why? Information that complements the data itself, including metadata, explanatory and technical notes, data quality statements and declarations enable transparency for both the data producer and data user. Assessments about data quality enable all parties to determine the fitness of purpose for the data, i.e. whether they can use the data for the purpose they had in mind.

Sometimes the fitness for purpose of the data won't be known prior to the data sharing being undertaken, for example where the data is being shared for analytical purposes. In these instances, this should be acknowledged in the agreement. Keep in mind that data that may be considered as being high quality for one purpose, may not be suitable for another purpose.

G. Data management

Data management should be an active part of the arrangement, and be an integral part of the data sharing agreement. Data management encompasses systems and processes to ensure data integrity and data security.

Data management practices to consider are:

- Metadata – use of authoritative or local standards, accessibility of metadata, version of the metadata relevant to the data being shared
- Data security – data transfer arrangements, data storage arrangements, data removal, data access requirements
- Access registers – maintenance of who has access to the data, when, duration and purpose, processes to update access lists for all parties, this may include signed copies of any documentation required for access to the data.

Why? Well defined and understood metadata can make data more useful. Other data management activities are good practice and may actually be required to meet legal restrictions.

In some cases, additional effort may be required to ensure a common understanding of the metadata and data security arrangements from both within the organisation and across organisations.

H. Costs

Document details of costs associated with the data sharing arrangement as negotiated and agreed by the parties. The costs may be associated with the transfer of the data, the extraction of the data, the data itself, access to the data, or the use of the data. A payment schedule may be included in the agreement.

Why? It is recognised that payment for data may be required due to business models and the costs associated with all aspects of data sharing. For instance, government agencies operating under the Australian Government Cost Recovery guidelines, or commercial organisations operating in the marketplace, may require payment.

WHAT MAKES A GOOD DATA SHARING AGREEMENT?

A good data sharing agreement should work well for all parties. A good data sharing agreement is:

- well-considered, documenting all relevant elements, components and considerations
- appropriate in language, length and level of detail
- signed by the appropriate personnel on behalf of all parties

and includes:

- clear statements of purpose, roles, responsibilities and expectations
- definitions of the data, including metadata, quality statements and outputs
- periodic reviews

CRITICAL SUCCESS FACTORS

The success of a data sharing arrangement goes beyond simply having an agreement in place. Some factors that can contribute to success are:

- a focus on desired outcomes
- ongoing regular communication
- a clear understanding of roles and responsibilities - important for building strong relationships
- flexibility - it may not be possible to agree on all of the details in advance
- time - allow enough time, it can take a significant amount of time to negotiate and develop a data sharing arrangement
- the identification of a champion within each agency to support the arrangement
- support of the arrangement at all levels within each agency, including executive, technical and operational levels
- a common understanding of the aims, purpose and context of the data sharing arrangement

It is important to note, a data sharing agreement supports the success of any data sharing arrangement.

REFERENCES

1. *Information Interoperability Framework: Sharing Information across boundaries*, Australian Government Information Management Office, Canberra, April 2006
2. *Statistical data integration across government: Driving information to support policy development and research*, Brian Pink, Australian Bureau of Statistics and Jane Halton, Department of Health and Ageing, PSM March 2009

FURTHER INFORMATION

For more about Metadata Management, see the [Keeping your data in good shape](#), and accompanying case studies.

For more about Data Quality, see the [ABS Data Quality Framework](#) (ABS cat. no. 1520.0).

For more about Creative Commons, see www.creativecommons.org.au

For more about the Government Information Licensing Framework, see www.gif.gov.au

For more information about Regional Spotlights, see www.regionalspotlights.com.au

For more information about the *National Government Information Sharing Strategy*, see, www.finance.gov.au/publications/national-government-information-sharing-strategy/

For further advice and assistance contact the National Statistical Service on inquiries@nss.gov.au

Please send your feedback and comments on this Guide to inquiries@nss.gov.au