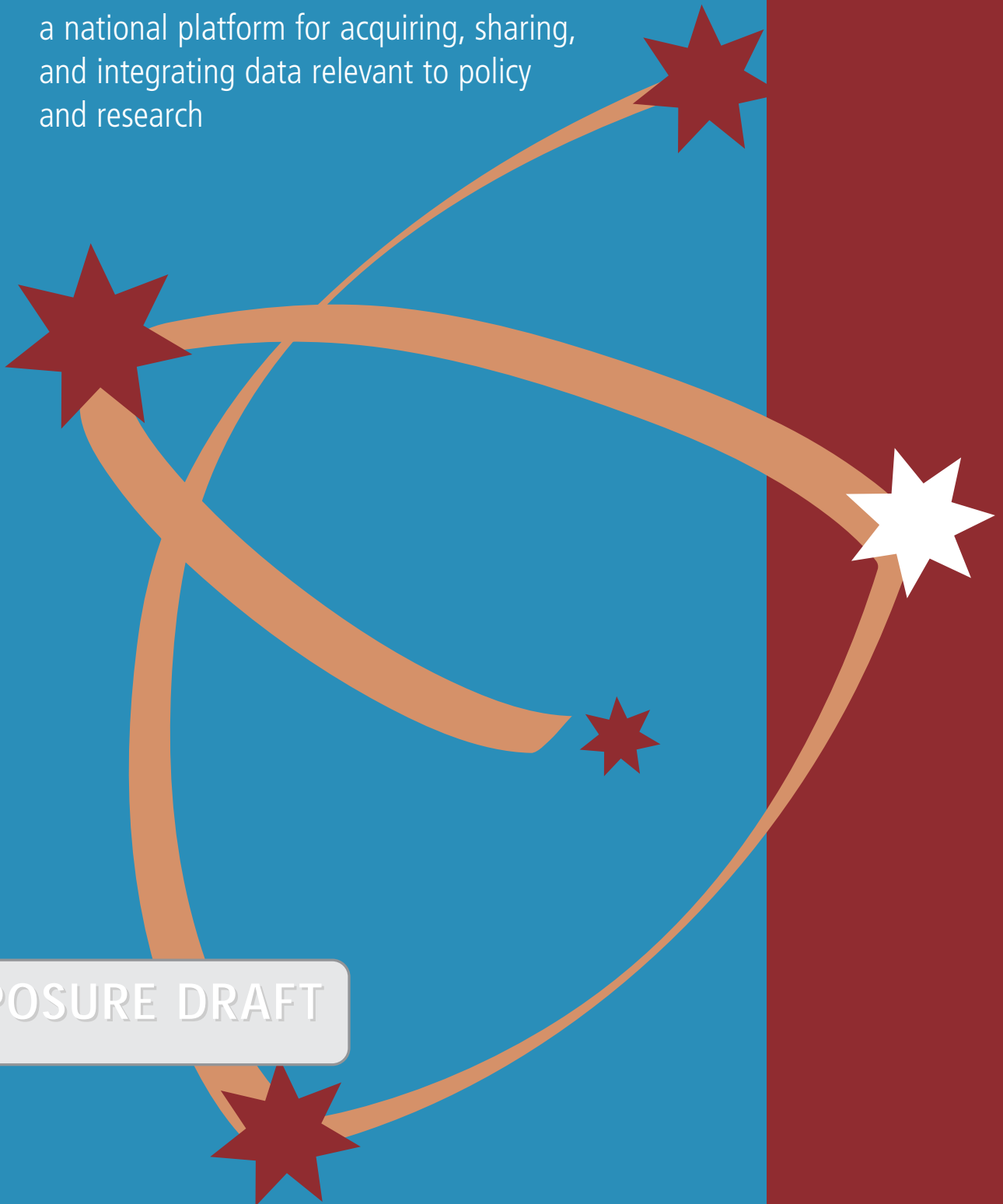




National Data Network

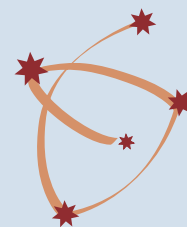
An Overview of the National Data Network

a national platform for acquiring, sharing, and integrating data relevant to policy and research



EXPOSURE DRAFT

TABLE OF CONTENTS



Introduction	2
Draft Guiding Principles	3
What is the National Data Network?	4
NDN Web Site	
NDN Central	
NDN Nodes	
Controlling Access to Published Data	
Publishing Data to the National Data Network	
Users of the National Data Network	
What are the Benefits?	7
Greater Collaboration between users and Custodians	
Access to Statistical Services	
Better Integration of Data	
Consistent Systems to Manage and Publish	
How will it work?	9
Node Software	
Metadata standards	
Assurance of Security	
Protection of Confidentiality and Privacy	
Interoperability with other networks	
Why is the Australian Bureau of Statistics involved?	11
How will it be progressed?	12
A Demonstration National Data Network	
Scope of the Demonstration National Data Network	
Potential Demonstration nodes	
Governance and the Interim Governing Board	
Who can participate?	14
Participating in the Demonstration National Data Network	
Following the Demonstration National Data Network	
Long Term Development Plan	
Appendices	16
1: Access Classes and Authentication Procedures	
2: Potential NDN Services	
3: Major Milestones in the Development	



Introduction

Australia has a rich collection of surveys, cohort studies, population databases and administrative records. These collections are held separately and are under-utilised in terms of their potential to yield new information of value.

The National Data Network will provide infrastructure, protocols, standards, and services to support the sharing and integration of data across Australia. It will be developed to increase the availability, accessibility, and usability of information sources relevant to policy analysis and research - particularly key administrative and survey data sets held by State and Federal Government agencies.

The National Data Network will provide data custodians with services, protocols, standards, and tools which will assist them to more efficiently manage and share data in a way that ensures security and privacy.

Users of the National Data Network will include agencies, policy analysts, researchers, and statisticians. The National Data Network will make the whole process of finding, understanding, acquiring, and making appropriate use of data easier and faster.

The National Data Network is being developed by the Australian Bureau of Statistics in collaboration with other interested agencies. The first stage will be the development of a Demonstration National Data Network during 2004 and 2005.

The purpose of this document is to describe key aspects of the proposed National Data Network to elicit feedback from stakeholders and expressions of interest from agencies wishing to participate in the development of the Demonstration National Data Network.



Australian Research Alliance
for Children & Youth

The creation of a "National Data Network" was proposed by the Australian Research Alliance for Children and Youth (ARACY). ARACY's vision is that Australia's rich collection of population surveys, cohort studies, population databases and administrative records be linked together to answer important questions about the wellbeing and life chances of Australia's children and young people.

In 2003 ARACY convened an "Expert Group" to develop an implementation strategy for the National Data Network. Workshops and Roundtable discussions were held in Canberra and Brisbane to gauge support and gather feedback and suggestions from a broad range of data custodians, researchers and data analysts. The National Data Network is one of several initiatives fostered by ARACY to promote collaborative research and the development of research infrastructure. Visit www.aracy.org.au for more information.



CSIRO has been a significant contributor to the vision and design of the National Data Network. CSIRO, through its work in the area of Health Data Integration, has developed deep expertise in the development of facilities for large scale, distributed information systems. This includes the development of protocols and facilities for secure data extraction, linkage, and analysis which preserve individual privacy and confidentiality. Visit www.csiro.au for more information.

Draft Guiding Principles



These draft Guiding Principles will represent a shared understanding about the use, development and operation of the National Data Network. The principles will be discussed, amended and adopted as required by the Interim Governing Board. Once agreed, they will guide future decisions made about the National Data Network.

Draft Guiding Principles

These principles recognise that official statistics and the data collected by surveys and held in the administrative registers of government agencies provide an indispensable element in the information system of a democratic society. This data provides governments, businesses and the people of Australia with information about the economic, demographic, social and environmental situation. In addition, they recognise that better reuse of data through the National Data Network and coordination among custodians is essential to achieve consistency, quality and efficiency in the Australian statistical system.

- 1.** Custodian agencies should maximise the amount of data which they expose through the NDN, so that:
 - Important statistics are compiled from key administrative and survey data sets relating to government programs and activities.
 - Ready access to quality statistics is provided to all Australians.
- 2.** Custodians should follow sound statistical practices for collecting, processing, storing and presenting statistical data before it is exposed to the NDN, including:
 - Using standard classifications, practices and frameworks and explaining deviations from the relevant international/national standards.
 - Where possible, ensuring that the statistical methodology and data remain internally consistent over time.
 - Allowing enough time to check the data for a reasonable level of accuracy and plan to release the statistics as soon as possible after their collection.
- 3.** Data which is exposed through the National Data Network should be explained by providing information about the methods and procedures used to collect and compile the statistics, so that users can interpret the data correctly. In particular, custodians should:
 - Provide analyses and explanations where they help the interpretation of statistics.
 - Be open about the quality of the statistics, so that users can better understand and interpret them.
 - Explain reasons for any changes that occur between collection periods.
- 4.** Custodians should invite and respond to comment on their data which is exposed through the National Data Network, including the erroneous interpretation and misuse of their data.
- 5.** An individual person's data which is published through the National Data Network should be used exclusively for statistical and research purposes and should not be used to identify individuals. Data exposed through the National Data Network should comply with privacy principles, confidentiality guarantees and other undertakings given to data providers.
- 6.** Researchers must only use the data ethically and adhere with undertakings given while accessing the data.

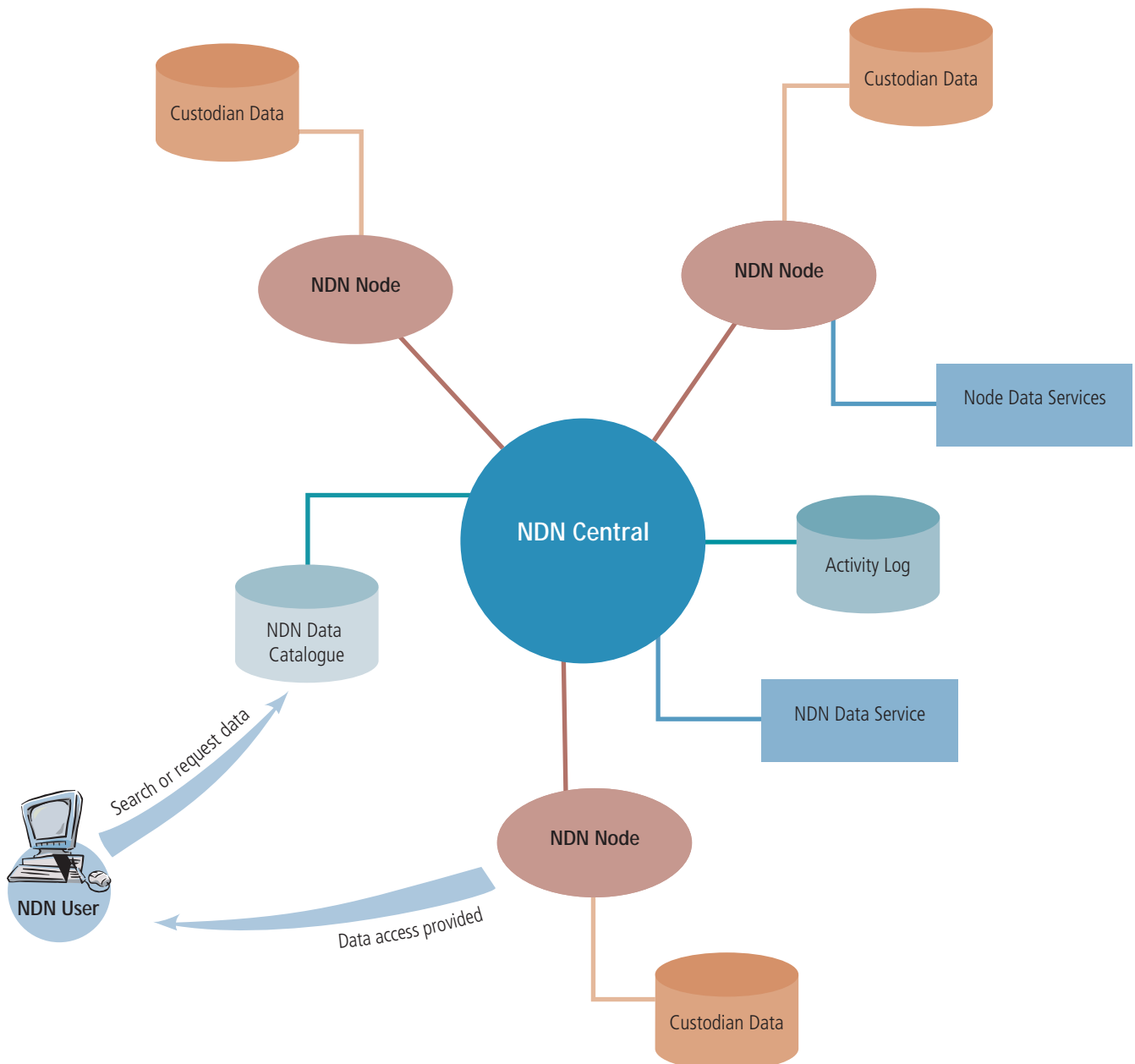


What is the National Data Network?

The National Data Network will create a distributed library of data holdings relevant to policy analysis and research. These data holdings will remain held and controlled by their Custodian organisations. Whilst data will be held by each Custodian, the National Data Network will provide a complete catalogue of available data sources to allow users to easily search for, and access, data holdings which have been published.

The National Data Network will consist of three key components, the NDN Web Site, NDN Central, and NDN Nodes. The NDN Central will be a hub in the network, linking the NDN Nodes and allowing entry by researchers and policy analysts.

Figure 1: Overview of the National Data Network



NDN Web Site

The NDN Web Site, <www.nationaldatanetwork.org>, will provide easy access to information about the network and link other relevant sites. The site will hold endorsed versions of guidelines, standards and protocols.

The NDN Web Site will allow feedback on directions and issues during the development phase. A range of discussion fora will be available to support collaboration amongst both users and developers. A series of white papers and draft documents will be made available through the web site for comment by stakeholders.

A range of administrative information will be stored on the site; such as statistics on usage, contact details and agendum, papers and minutes of National Data Network committees and boards.

Following the development stage, both documentation about the system and software downloads will be available from the site.

NDN Central

NDN Central will act as the hub of the National Data Network. Each set of data or data service available through the National Data Network will have a unique resource description stored in a consistent format. The central server will compile descriptors of the available data from the Custodian's web site.

NDN Central will maintain a central metadata repository, which will contain descriptions of all available NDN data sources and services. NDN Central will include a search facility to allow users to search and retrieve relevant information about available data based on the metadata repository.

NDN Central will also provide some general statistical services. Typically, these services will be commonly accessed services which require little maintenance. NDN Central will also log activity conducted through the network to ensure that it is used appropriately.

NDN Nodes

Custodian agencies will be connected to the NDN Central via the World Wide Web and will operate as NDN Nodes.

An NDN Node is a library of metadata about the data which the custodian has published to the National Data Network. Custodians will have complete control over the information contained in their node.

Each node will consist of a base program which contains information about where data is stored and key metadata elements about that data. Some NDN Nodes may wish to expand their role by operating selected statistical services. These statistical services would also reside on the NDN Node.

When the National Data Network is established, all government bodies will be able to utilise the National Data Network software. The only requirement to participate is a commitment to the National Data Network Guiding Principles. While NDN Nodes will typically be federal or state government institutions, other reputable organisations with quality data holdings relevant to policy and research may be included as NDN Nodes.

Publishing Data to the National Data Network

Custodians will be able to select a range of their data holdings and publish it through the National Data Network. Any data may be loaded to the network, so long as it is accompanied by an appropriate “Data Resource Description” (or metadata). This metadata can be used by policy analysts and researchers to decide which data is relevant for their purposes and assist them to interpret and analyse that data. The metadata will also describe and control access and use restrictions that might apply to the data.

Both aggregated data and unit level data will be able to be published through the National Data Network.

Controlling Access to Published Data

Custodians will control access to their data holdings. To safeguard data, a range of standard Access and Use Policies will be developed to provide appropriate levels of security for varying types of data. Four basic levels of security are envisaged - unrestricted, approval required, remote access only and specification only. Further information about these access classes is given in Appendix 1.

Custodians will be required to determine appropriate levels of security and attach an appropriate access policy to each data source. The National Data Network will ensure that the access rules and processes stipulated by the Custodian are rigorously applied.

Access to data will be able to be restricted to groups of users. These groups may consist at various levels, such as users within:

- the custodian agency;
- a specified agency;
- a jurisdiction, such as a state; or
- a field or collection of agencies, such as agencies involved in delivering health services.

If required, custodians will also be able to develop and integrate specific security features into the network.

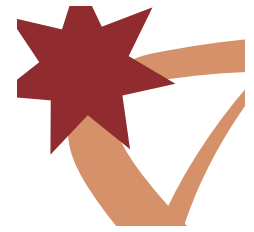
An example: Copies of ABS Confidentialised Unit Record Files (CURFS) are only provided to users who have signed a specific “Undertaking” and the files are delivered in encrypted form. The National Data Network will support users by making these access conditions clear and by helping the user complete and submit the appropriate form to the ABS. When the ABS receives a request for CURF data, it will be advised by the National Data Network whether the appropriate conditions have been met (ie. an Undertaking submitted and accepted) and the data will be automatically encrypted before transmission.

Users of the National Data Network

The National Data Network will service all of Australian society. The users will range from policy analysts from the government sector, through to academic researchers and private sector. These users will be able to search across the vast data libraries which form the network, access selected data, utilise integration services and republish to the network.

Users will only be able to utilise the network after they have been appropriately authenticated.

What are the Benefits?



The National Data Network will, over time, deliver significant value and benefits to:

- agencies, by avoiding development costs;
- policy makers, by supporting efficient access to data;
- researchers, by articulating clear and consistent access protocols; and
- the community, by improving outcomes from government policies.

The National Data Network will facilitate greater collaboration, faster access to data, improved integration and consistent systems across government.

Greater Collaboration between users and Custodians

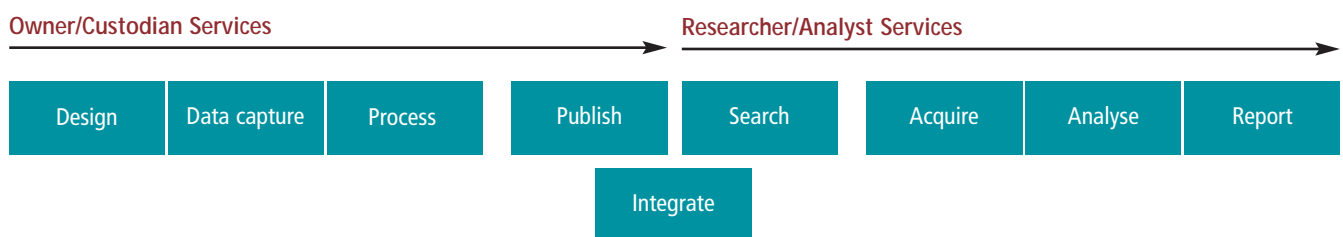
The National Data Network will facilitate greater collaboration between users and Custodians. Knowledge will be transferred between users and custodians through provision of collaborative work and discussion spaces. These environments will help people share experiences and lessons learned.

Access to Statistical Services

The National Data Network will provide access to a range of services to support the creation, management, integration and analysis of data. Some of these services will be provided through the NDN Central, while others will be provided by third party service providers. These third party providers will only be able to register their services with the approval of the governing board.

Progressively, the National Data Network services will be expanded to cover the full life-cycle of statistical processes. Services will be provided to aid Custodians in designing, creating and disseminating statistical data; and researchers in analysing this data.

Figure 2: Services Framework



Appendix 2 provides examples of potential services offered through the National Data Network.

Better Integration of Data

Complex issues often require the integration of data from multiple sources. For example, a study of factors leading to higher student retention levels might require integration of data about educational outcomes with income and housing and health data.

Integration of data sets can be achieved at either the aggregate level or the individual “unit” level. Many purposes can be satisfied through linkage of data at an aggregated level, for example aggregate cells from different datasets may be integrated by a common geography. Other applications may require integration at the unit record level, such as linking all health events for an individual.

The National Data Network will promote stronger integration of data by providing access to metadata, standards and guidelines (classifications, data elements, question formats, concept definitions, statistical frameworks). These standards will promote the design of data collections that are compatible with other sources so that the value of these data sets can be augmented.

The National Data Network will support integration of data in ways that meet custodian obligations to maintain confidentiality, protect intellectual property and comply with privacy principles. It is intended that the National Data Network will offer services to integrate data at both the aggregate and unit levels. These services will be designed to avoid disclosure of identifying information.

Consistent Systems to Manage and Publish

The National Data Network software can be used for both managing data within an agency and publishing this data through the network. Utilising a single system to undertake both functions can reduce inconsistencies and double handling.

How will it work?

NDN Node Software

Custodians will be connected to the National Data Network using “NDN Node Software”. This software is being developed using an Open Source tool kit which will allow Custodians to download the software free of cost and allow inspection and contribution by developers. To participate as a node of the National Data Network, a Custodian must be approved by the NDN Governing Board. In line with the NDN Guiding Principles, Custodians who utilise NDN Node Software will be required to:

- publish data holdings relevant to policy and research;
- maintain metadata which is consistent with the current NDN Metadata Standards;
- operate NDN Node Software configured in line with the NDN Node Configuration Guide; and
- participate and cooperate with any audits which have been determined by the Board.

Metadata standards

Metadata will be stored as a NDN Data Resource Description which describes the data and how it can be accessed. National Data Network metadata will reflect international standards and other metadata standards used by the Australian Government. NDN Data Resource Descriptions will document key aspects of the data and will be the base for searches across the network. This metadata will be compiled in the central NDN Metadata Repository and regularly updated.

Metadata standards for the National Data Network will be developed in consultation with other key stakeholders. Two descriptive metadata standards are currently in use in the government sector, specifically:

- ANZLIC metadata standard, which is specific to geospatial data; and
- Australian Government Locator Service, which has a broader focus.

These two standards are compatible with the Dublin Core Metadata, which is the base for international standards for metadata. National Data Network metadata will be required to align with both these standards and will be developed in consultation with the users and developers of ANZLIC and Australian Government Locator Service.

Assurance of Security

Authentication is the process of verifying that the person or organisation at the other end of an on-line transaction is who they are expected to be.

The National Data Network will make Authentication processes visible to Custodians so that they can understand and rely on the approach taken in relation to their own data holdings.

The authentication strategy for the National Data Network will be consistent with the Australian Government Authentication Framework (refer to Appendix 1 for more detail). This framework identifies four levels of risk and maps suggests appropriate authentication techniques to each level. Appendix 1 provides examples of how National Data Network transactions might map to each risk level. For the majority of transactions on the National Data Network authentication will be via userids and passwords. It is envisaged that most users will register with the National Data Network by lodging (on-line) a request with their personal details. After these details have been checked, probably by referring the request to a trusted agency, the user will be issued with a userid and password.

For high risk datasets, Custodians may require a stronger authentication in the form of a digital certificate issued after a rigorous check of identity. The National Data Network will accept such credentials from approved Registration Authorities such as the Health eSignature Authority.

The National Data Network will monitor the progress of access protocol, such as the submission of legal undertakings. The network will ensure that data is not released until all appropriate undertakings have been signed. By codifying and automating common data access protocols, the National Data Network will make data supply and access easier, faster and cheaper for both custodians and users.

Protection of Confidentiality and Privacy

The National Data Network will assist Custodians in implementing rules and procedures which maintain the confidentiality and privacy of their data providers. Activities such as data integration, may require the application of confidentialisation techniques to reduce the risk of disclosure of sensitive information.

Confidentialisation and disclosure avoidance tools will be available to mask or collapse the cells in aggregate data which are most likely to lead to identification of individuals data. These tools will support a range of techniques and will be able to be adjusted to align with existing standards maintained by Custodians.

The network will also provide services which facilitate analysis without data leaving the Custodian's node. Initially, these services will be based on the Australian Bureau of Statistics' Remote Access Data Laboratory (RADL).

The National Data Network will provide automated work flows to support processes such as ethics committee submissions related to access and use of confidential data. In many sectors, ethics committees exist to assess whether protocols and techniques proposed by researchers will preserve privacy. The National Data Network will allow researchers and Custodians to access templates for submissions and relevant text explaining techniques embedded in the network.

Interoperability with Other Networks

The National Data Network will interoperate with existing networks and facilities which play a role in the sharing of Australia's information assets. The NDN Metadata Standards will be designed to promote interoperability with other networks, in particular, with Australian Spatial Data Infrastructure. The Australian Spatial Data Infrastructure provides access to more than 39,000 spatial datasets held across 26 nodes.

Why is the Australian Bureau of Statistics involved?

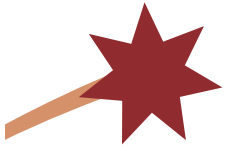


The Australian Bureau of Statistics considers that the National Data Network will strengthen Australia's statistical system. The Australian Bureau of Statistics is leading the National Data Network as a key element to grow the National Statistical Service. The Australian Bureau of Statistics' objectives for the National Statistical Service include:

- a shared commitment at both Commonwealth and State government levels to the concept of a national statistical service that delivers the statistics required by key users no matter what their source;
- active participation in the development of standards for electronic information exchange and underlying data sets (e.g. accounting standards); and
- better utilisation of both public and private administrative and transactional data sources for national statistical service purposes.

The Australian Bureau of Statistics recognises the potential of the network in building stronger coordination of Australia's statistical activities. The *Australian Bureau of Statistics Act, 1975* identifies a function of the Australian Bureau of Statistics to "ensure coordination of the operations of official bodies in the collection, compilation and dissemination of statistics and related information". The Australian Bureau of Statistics is supporting the National Data Network because of the benefits to all of Australia's official bodies in both collecting statistics and coordinating data sources.

The National Data Network will both assist the Australian Bureau of Statistics in achieving its function and provide a low cost suite of statistical tools to aid all governments agencies.



How will it be progressed?

A Demonstration National Data Network

Implementation of the National Data Network will commence in 2004 with the development of a "Demonstration NDN" led by the ABS. The Demonstration NDN will showcase the functionality of a National Data Network.

The specific outcomes of this project will be:

- to identify the feasibility of developing a networked system (utilising open source software) to share data and provide on-line statistical services in the government context;
- to understand the cost to Custodians of implementing and participating in a National Data Network; and
- to understand the issues associated with the sharing of secure information in accordance with the policy of each agency.

A small consortium of agencies is being established to develop the Demonstration NDN. This consortium will establish a core system underpinned by sound statistical protocols.

Scope of the Demonstration NDN

The Demonstration NDN will expose a selection of relevant data sources and show the benefits of the core system. The data will be sourced from the data holdings of the consortium and will be loaded and documented using the suite of tools offered through the system.

It is anticipated that the Demonstration NDN will include a range of services. The core system must include facilities to establish nodes in each agency, document the data holding, search across data holdings and facilitate access to the data. Additional services are being considered for inclusion, such as tools to improve confidentialisation of data. The services which will be included in the Demonstration NDN will be finalised by the interim governance arrangements.

Maintaining the security of data exposed through the system will be a priority. The consortium will explore a range of security requirements and the associated cost of embedding solutions into the systems. Many of these solutions will involve integrating existing security facilities. Each security option will be assessed for its ability to manage access to sensitive data which is consistent with custodians obligations.

Potential Demonstration nodes

The Australian Bureau of Statistics has held discussions with a wide range of State and Australian Government agencies and several agencies have expressed a strong interest in participation in the demonstration network. Demonstration nodes will be announced during October 2004.

Governance and the Interim Governing Board

The Interim Governing Board will be convened by the Australian Statistician to govern the development of the National Data Network (until more permanent governance arrangements are established during 2005).

Membership of the Interim Governance Board requires participation in the Demonstration NDN systems development, protocols or exposure of data to the National Data Network. This participation involves a contribution to any number of the programs or procedures governing the operation of the National Data Network.

The Interim Governance Board will be responsible for determining the directions of the National Data Network in relation to:

- principles underlying the National Data Network;
- priorities relating to the development of National Data Network initiatives and projects;
- terms of reference associated with National Data Network projects;
- protocols for the long term governance of the National Data Network; and
- development of communication strategies across members of the National Data Network; incorporation of the views and needs of stakeholders within the directions of the National Data Network.



Who can Participate?

Participating in the Demonstration NDN

The development consortium for the Demonstration NDN will be established during October 2004. It is expected that the consortium will consist of approximately twelve government agencies from both State and Australian Governments. These agencies will contribute in different ways to the development. Some agencies may choose to:

- publish data to the Demonstration NDN;
- contribute to the development of standards, such as for metadata;
- develop additional statistical services to link into the network;
- provide administrative support for aspects of the network; or
- guide the development by participating on the Interim Governance Board.

Invitations to participate in the development consortium will be sent during September 2004. Interested agencies will be asked to sign a "Statement of Intent" which articulates their expected contribution to the network.

Other interested agencies should contact the NDN Administrator to identify ways in which they can participate.

Following the Demonstration NDN

Throughout the development, the consortium will seek stakeholder feedback. Lessons learnt while applying the protocols and establishing the systems will be distributed for comment. A regular newsletter and the National Data Network web site will be used to obtain feedback from stakeholders. Feedback will be considered through the interim governance arrangements before the demonstration system is finalised.

The Demonstration NDN will be released in April 2005. At this time, the consortium will demonstrate the system through a series of on-site presentations. These presentations will display the functionality, utility and flexibility of the system and explain the cost and issues with implementation.

A stakeholder conference will be held in the middle of 2005 to identify future directions for the National Data Network. Issues determined will include permanent governance arrangements, alterations to the core systems and priorities for future development.

Long-term Development Plan

The development of the National Data Networks full functionality will be progressed incrementally. The first stage will be the development of the Demonstration NDN. Following stages will release progressively more functionality. The following diagram presents an indicative suite of functionalities will be released at each stage. More detail about these functionalities is given in Appendix 2.

Figure 3: Indicative Release Program for the National Data Network.

Stage 1	Stage 2	Stage 3
Demonstration NDN	Release Version 1	Release Version 2
NDN Central	Confidentiality tools	Automatic Coding
NDN Node Software	Coding Assistant	Character Recognition
Metadata tools	Data item Definitions	Collection Control
Access Policy Tools	Standard Questions	Sample Selection
Search Facilities	Estimation Tools	Forms Design
Data Requests	Aggregation Tools	Archiving
Data Acquisition	Data Matching	Seasonal Adjustment
Authentication	Supervised Analysis	Imputation
Security	Subscription Service	Analysis and Graphing
Data Integration		Research Repository
Remote Access		
12 - 18 Months	18- 36 Months	36 - 72 Months



Appendix 1

Access Classes and Authentication Procedures

1.1 Access Classes for the National Data Network

The authentication infrastructure will support the access protocols and security requirements specified by the Custodian. These access classes will be: unrestricted, approval required, remote analysis only and specification only.

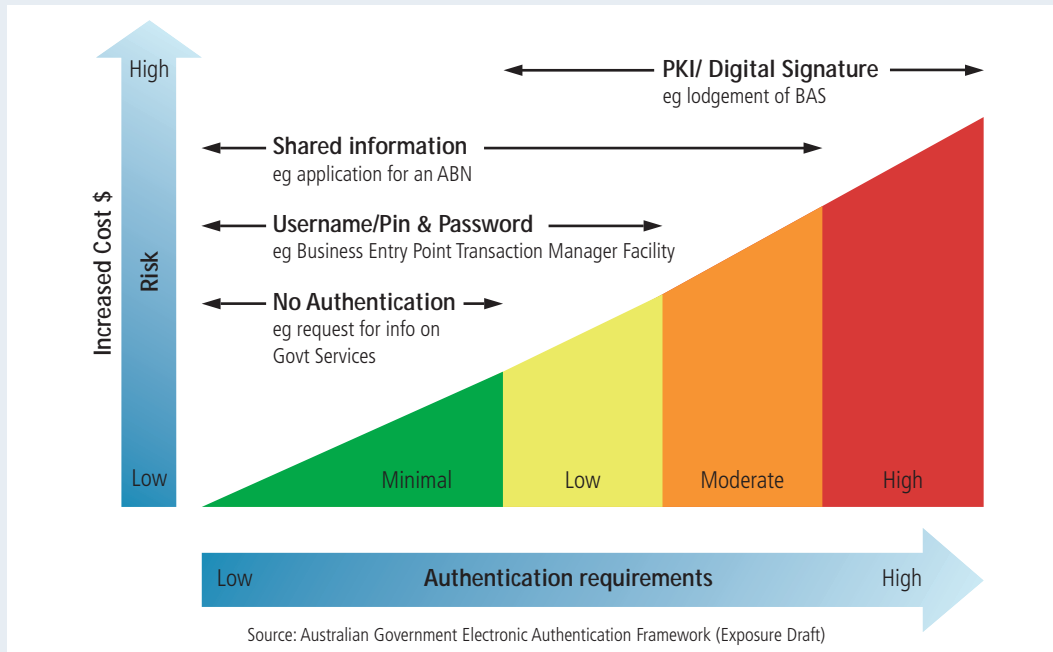
- 1 Unrestricted access** - users can freely access the data object, metadata or service without restriction. This type of access might be applied to the access of unsecured publications, material or data cubes that is currently provided by the Custodian to the full range of users.
- 2 Approval required** - individual users or groups must apply for access to data objects or services through a specified access protocol. This protocol is defined by the Custodian. These access protocols will include at least one of the authentication classes indicated in 1.2, particularly password authentication procedures or the use of access tokens (PKI). Each user will need to apply for the appropriate certification either through the registration authority or by directly consulting with the Custodian agency. Direct access to the data object will typically be granted through the National Data Network system once the appropriate authority has been granted. Common examples of this type of access level is when the data object is subject to a charge by the Custodian or that an undertaking needs to be completed by the potential user before access will be granted.
- 3 Remote Analysis Only** - individual users will not acquire direct access to the data object. Custodians will again specify strict access protocols. Users can submit programs that will be applied to data and the results returned following the signing of undertakings specified by the Custodian access protocol. The release of this information is likely to require strict authentication procedures. The results of 'remote access only' requests will be vetted by the Custodian prior to release in accordance with current agency release procedures.
- 4 Specification Only** - some data objects referenced within the National Data Network will not be available for manipulation by non Custodian agencies. These data objects will lie in the 'specification only' access class. In this class, users submit data specifications indicating cross-tabulation and output requirements to the Custodian agency. The Custodian agency prepares the output as per these specifications (where practicable) and applies the standard release procedures of the Custodian agency. These data objects are likely to be subject to the highest access and authentication procedures.

1.2 Authentication procedures for the National Data Network

The Australian Government Authentication Framework (AGAF) seeks to provide guidance for the development of authentication procedures associated with Australian Government on-line transactions.

Broadly, each of the 4 risk levels described in Diagram 1.1 are primarily associated with the level of risk ascribed to the possibility that a false assertion of the user's identity has been made. As the consequences of such a false assertion increase, so too does the level of authentication required. A more detailed outline of the elements of these risk categories can be found at the Australian Government Information Management Office <www.agimo.gov.au/infrastructure/authentication>.

Diagram 1.1: AGAF Range of Authentication Requirements and techniques



In line with this framework, the Custodian will ascribe a risk level and access requirements to each individual transaction type. The National Data Network will then apply the appropriate authentication and access protocol. The network is likely to use a combination of password authentication and digital signature procedures. Where possible the network will use current e-signatures for authentication. Diagram 1.2 includes examples of how National Data Network transactions might map to each risk level indicating the range of authentication procedures required.

Diagram 1.2: NDN transactions and associated risk levels

Authorisation requirements		National Data Network transactions
Level 1 Minimal risk	No requirement for confidence in the assertion	Browsing of 'public' metadata, data and services Lodgement of feedback, requests for information
Level 2 Low risk	Some confidence in the assertion	Participation in discussion/collaboration forums
Level 3 Moderate risk	Moderate confidence in the assertion	Access to some 'restricted-access' data
Level 4 High risk	High confidence in the assertion	Proof of (significant) payment Access to sensitive data

Appendix 2 Potential NDN Services

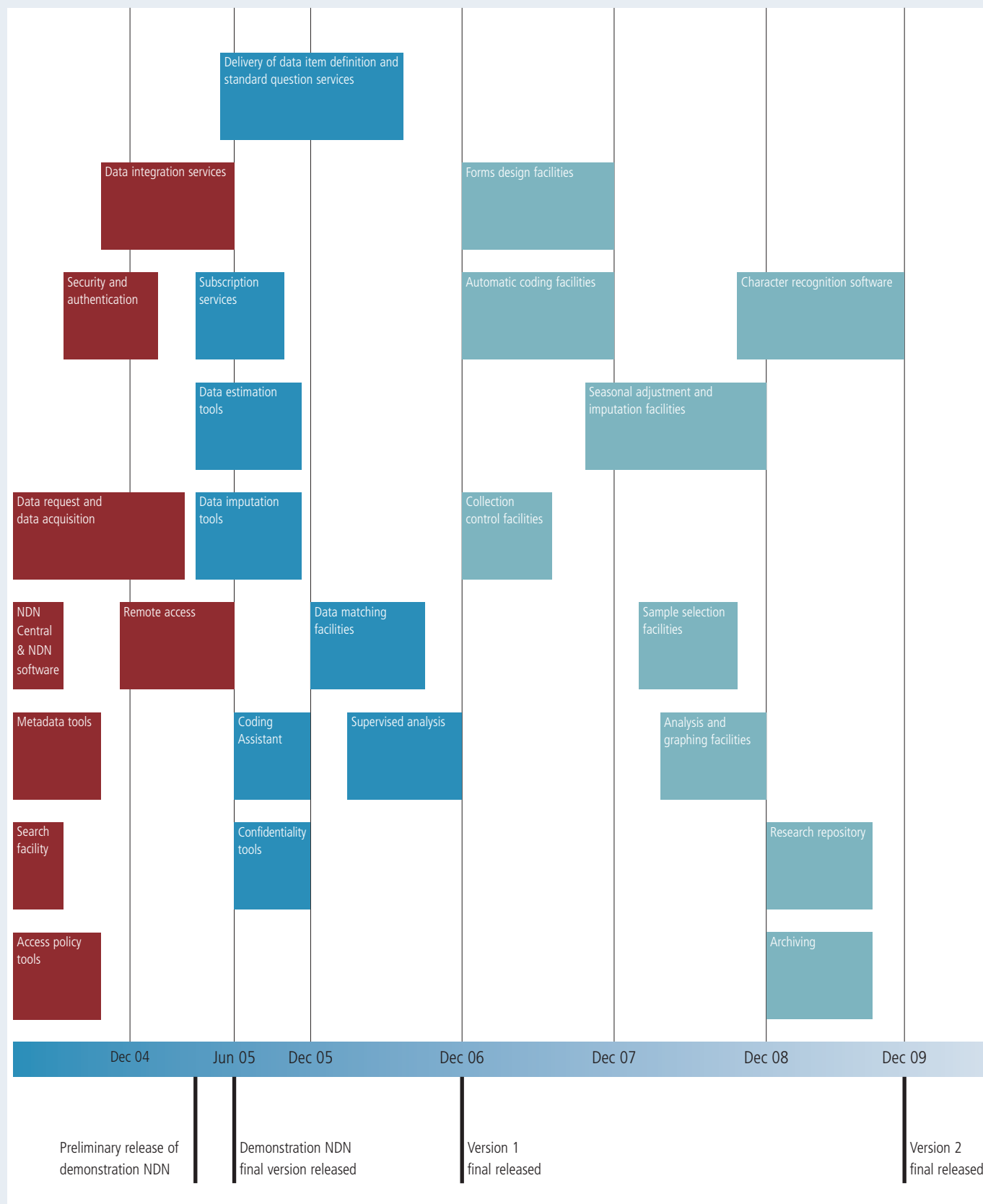
The National Data Network will provide access to a range of services designed to support the creation, acquisition and use of data resources. The following framework has been used to cluster/ classify services depending on where they fit in the data “life cycle”:

This table provides more detail on services which might be provided by the NDN:

Design	Design standards	Registry services, access to standard data Item definitions, classifications, questions modules, derivation specifications
	Sample selection	Search tool to source possible population frames, sample size calculator, sample selection tool
	Forms design	Form design principles, design templates and services
Data Capture	Coding facilities	Automated and assisted coding services
	Character recognition	OCR software, standards and services
	Collection control	Collection management services
Process	Edit	Edit specifications, edit services, edit strategies
	Estimation and imputation	Seasonal adjustment, estimation and imputation services
Publish	Confidentialise	Confidentiality protocols and standards services for confidentialisation of data
	Document	Tools and services for documentation of data and services, documentation standards
	Expose Archive	Publish data and service resource definitions to the NDN Data archive, backup and recovery services
Link/Integrate	Match/Link	Data matching and linking services
Search	Search/Request	Search all data collections which have been published to the NDN, request access to data and processing services
Acquire	Acquire test data	Provision of data for testing/ evaluation purposes
	Process access Request	Processing of access requests as per custodian defined access and use provisions
	Extract Subscribe	Extraction of a subset of data from a database Regular, automated, delivery, triggered by an update, correction or new release
Analyse	Supervised analysis	Remote analysis service which allows users to submit programs which are then run by the custodian agency (program code and results are usually vetted).
	Analyse/graph/tabulate	Analysis, tabulation and graphical presentation services.
Report	Document data source issues	Minimum quality description for data reporting, sourced from NDN documentation
	Expose findings	Provides for the publishing of secondary data sources through the National Data Network

Appendix 3

Major Milestones in the Development



© Australian Government 2004

A complete list of publications is available
at our website:

www.nationaldatanetwork.org

Any enquiries about or comments on this
publication should be directed to:

inquiries@nationaldatanetwork.org

September 2004

