NATIONAL ARCHIVES OF AUSTRALIA

Interoperability development phases resource

January, 2019

Australian Government
National Archives of Australia

Table of contents

Introduction 5

1.1. Building interoperability 5

Why interoperability? 5

Implementation advice 6

Other interoperability initiatives 6

2.1. Interoperability key themes 8

Business 8

Security 8

Legal 9

Semantic 9

Technical 9

3.1. Interoperability development phases 11

4.1. Data governance and management 13

Why data governance and management? 14

Data disposal 14

Data champion 14

4.1.1. Metadata for interoperability 15

Metadata strategy 15

Metadata harvesting 16

Metadata repository 16

Metadata statements 16

Metadata for publication and exchange 17

Common architectural approaches 17

4.1.2. Data indexing and discovery 18

Data profiling 19

Explore your agency's data 19

Search tools 20

4.1.3. Taxonomy 20

Taxonomy creation 21

Automation tools 21

Ontology 21

Languages 21

4.1.4. Open data and formats 22
Open file formats 22
Examples of common open file formats 22
Open data 23
4.1.5. Data quality 24
Data quality assessment 24
Data quality tools 25
Data quality and metadata 25
4.1.6. Data compliance and security 26
Secure data exchange 27
Privacy and de-identification 27
Licensing and terms of use 28
5.1. Current state assessment 29
Features and tasks 30
6.1. Future state vision 31
7.1. Gap analysis 34
8.1. Planning and design 36
Features and tasks 36
9.1. Implementation 39
9.1.1. Data integration 40
Data integration approaches 41
Web data services 42
Application Programming Interfaces (APIs) 42
API management 42
Useful guidance on developing APIs 43
Architecture models 43
Extract, Transform, Load (ETL) technologies 43
9.1.2. Data collection 44
9.1.3. Data migration 46
Understand your data and its quality 46
Identify stakeholders 46
Data extraction 46
Data remediation – improving the quality of data 47
Data migration techniques 47
Data validation, auditing and verification 47

9.1.4. Data exchange 49
Exchange standards 49
Exchange agreements 49
Data exchange languages 49
Platforms for data exchange 49
9.1.5. Business As Usual (BAU), monitoring and improvement 50
Monitoring and continuous improvement 50

Introduction

This learning resource helps build your understanding of interoperability by taking you through six core development phases that address some of the activities you would need to do to build your agency's interoperability. The resource does not intend to replace your agency's preferred project management methodology – and it does not cover all possible interoperability approaches, tools and technology.

You should read this resource with the interoperability scenarios document.

1.1. Building interoperability

The Digital Continuity 2020 Policy calls for information, systems and processes to be interoperable by 31 December 2020.

The potential of your agency's information and data to be shared and re-used depends on its quality, discoverability, and overarching management.

To achieve this potential, your agency needs to build interoperability across a range of themes - business, security, legal, semantic and technical.

Building interoperability means you can exchange information and data between different systems. It is dependent on clear, shared expectations for the context and meaning of data across systems.

Data governance and management is an essential component of information governance, particularly in data-centric agencies and is the foundation for building interoperability. It supports standardisation and management to facilitate discoverability and sharing.

Why interoperability?

Interoperability supports the use and reuse of government information and data as key assets and can:

- provide consistent, coordinated and more timely services
- improve accessibility
- lessen the impact of structural changes in government
- reduce the risks of technical obsolescence
- inform policy development and decision-making
- reduce the cost of information and data management through reuse and shared infrastructure.

Implementation advice

The following implementation advice helps build your agency's capability to meet Digital Continuity 2020 Policy interoperability expectations:

Principle 3 Information is interoperable	Implementation advice
Agencies will have **interoperable information, systems and processes** that meet standards for short and long-term management, improve information quality and enable information to be found, managed, shared and re-used easily and efficiently.	Interoperability key themes help you understand how interoperability is not just a technical fix, as it also relies on working with your information and data to align your business, security, legal and semantic needs. Interoperability development phases will help you plan and implement solutions to address interoperability hurdles that are visualised in the interoperability scenarios. Your results from using the Business System Assessment Framework (BSAF) can be used to identify: • the need to *integrate* business systems or to *migrate/export* data to address risks or gaps • system functionality to meet your information and data needs over time • what information and data is held in your systems and its value. Minimum metadata supports a standards based approach to sharing information and data.

Other interoperability initiatives

Other interoperability initiatives are being led by a range of Australian Government agencies and include the:

- Office of the National Data Commissioner framework for data sharing and release
- Digital Transformation Agency agenda covers a range of programs including whole of Federal Government digital Platforms and a Secure cloud strategy
- Department of the Prime Minister and Cabinet is leading the Data integration partnership for Australia (DIPA) and provides Guidance on data sharing for Australian Government Entities
- Multi-agency data integration project (MADIP)

Related links (naa.gov.au)

- Business System Assessment Framework
- Digital Continuity 2020: Agency implementation targets and pathways poster.
- Digital Continuity 2020: Agency implementation and pathways
- Information governance
- Interoperability: Data governance and management
- Interoperability scenarios pdf
- Minimum metadata set

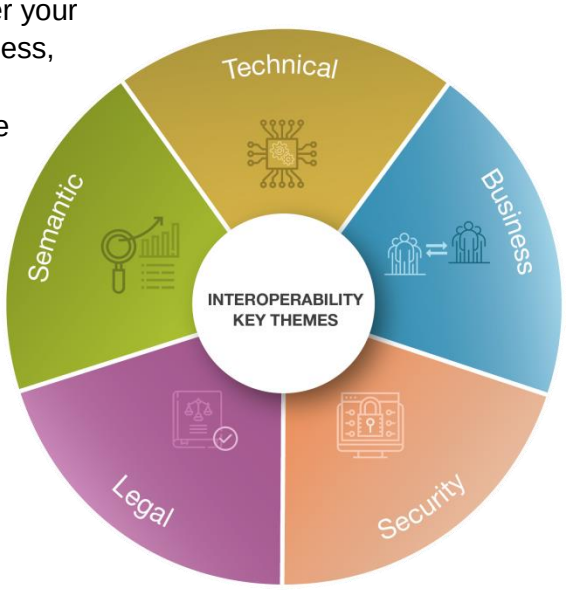
Related information

- Data Integration Partnership for Australia (DIPA) – Department of the Prime Minister and Cabinet (PM&C)
- Data Interoperability Standards Consortium - (DISC; United States)
- Cloud strategy - Digital Transformation Agency (DTA)
- Guidance on Data Sharing for Australian Government Entities – PM&C
- Multi-Agency Data Integration Project - Australian Bureau of Statistics (ABS)
- The Open Data Institute - (ODI; Australian network)
- Project interoperability - (ODNI; United States)

2.1. Interoperability key themes

Interoperability is built around: **business**, **security**, **legal**, **semantic** and **technical** themes. Building interoperability needs technical solutions; however your success and maturity will vary according to how well business, security, legal and semantic considerations have been implemented. The key themes address the complex nature of creating and maintaining interoperable information, systems and processes.

Your information governance committee or equivalent can help you coordinate initiatives across the themes by implementing information and data governance mechanisms, such as frameworks, policies and standards.



The business theme ensures processes, responsibilities and expectations for exchanging data and information are aligned across all relevant stakeholders.

Key considerations include:

- understanding how information and data assets are used to meet business outcomes
- working with stakeholder and community and expectations eg, designing and delivering services for stakeholders that span across Government agencies
- understanding operational and strategic requirements and expectations for sharing and receipt
- aligning requirements to similar enterprises (for example financial transaction regulation, geospatial and criminal intelligence)
- implementing an 'open by default' position and identifying exceptions through risk assessments.

The security theme ensures data is protected and the risks of data and information exchange are managed.

Key considerations include:

- determining security classifications and user permissions
- meeting Protective Security Policy Framework and Information Security Manual requirements, eg, for the authentication and encryption of data
- working collaboratively to develop comprehensive risk assessment procedures for information and data
- undertaking data risk assessments across multiple business systems including those that target concerns such as sharing sensitive information

- developing standard workflows for removing personal or identifiable information from data to be published
- verifying your published datasets on a routine basis.

The legal theme ensures legal, privacy and ethical requirements are met when exchanging or using information and data.

Key considerations include:

- meeting legislative, regulatory and ethical requirements including Australian Privacy Principles
- ensuring source and combined data is trusted and complete
- determining custody, ownership, sharing or usage rights
- creating data sharing arrangements, including memorandums of understanding, service level agreements or licences.

The semantic theme ensures that users and systems understand the meaning of exchanged information and data as intended.

Key considerations include:

- tools to control the use of terms and language, including taxonomies, thesauri, ontologies, data dictionaries and classification schemes
- using metadata standards and schemas
- using data quality reporting outputs to help understand and assess the usefulness of data
- developing and maintaining enterprise data models, master data and data quality standards.

The technical theme enables platforms, systems and applications to exchange or process information and data.

Key considerations include:

- using whole-of-Government platforms
- selecting machine readable file formats – and where feasible, open and not 'locked in'
- using Application Programming Interfaces (APIs) to allow consumers to integrate with agencies' data
- publishing web services to enable the sharing of data
- using tools to help catalogue and profile information and data across an agency
- creating digital services such as customer portals which streamline data collection and validation
- using technologies to transform and improve legacy data
- using standardised data exchange specifications.

Related links (naa.gov.au)

- Building interoperability
- Classifying information
- Information governance
- Information governance committee
- Interoperability: Data governance and management
- Interoperability development phases
- Interoperability scenarios pdf

Related information

- Australian Privacy Principles - Office of the Australian Information Commissioner (OAIC)
- European Interoperability Framework: Implementation Strategy - European Commission
- Data Interoperability Standards Consortium - (DISC; United States)
- Digital Service Standard - Digital Transformation Agency (DTA)
- Information Security Manual - Australian Signals Directorate (ASD)
- Open Data – Department of the Prime Minister and Cabinet (PM&C)
- The Open Data Institute - (ODI; Australian network)
- Platforms for Government - DTA
- Protective Security Policy Framework - Attorney-General's Department (AGD)
- Project interoperability - (ODNI; United States)

3.1. Interoperability development phases

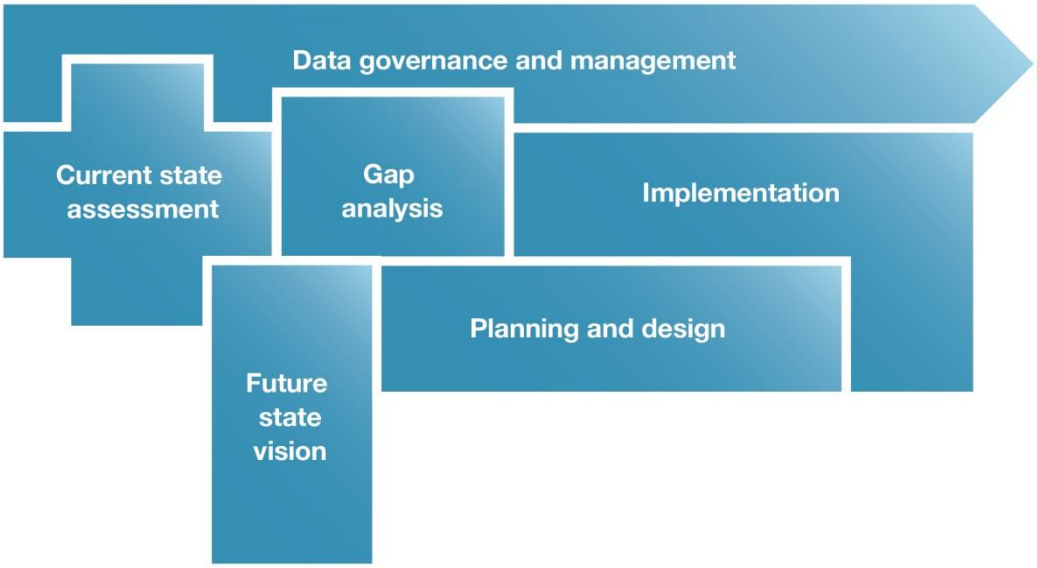
We've created a learning resource that covers six core phases to help you build interoperability in your agency.

Information in the six phases will help you:

- learn about typical activities to resolve four common interoperability hurdles outlined in the interoperability scenarios
- understand how to approach interoperability projects including who needs to be involved and learn about common interoperability technologies
- understand the importance and role of governance and management
- plan your own interoperability projects with the available skills and resources. Eg, with limited resources you can make a start with data governance activities and conduct a gap analysis based on your current and future state.

The development phases do not:

- replace your agency's project management methodology
- provide step-by-step instructions. In an agile environment you can use them with other actions, omit them or operate them in parallel.



Data governance and management ensures your data is standardised and managed to ensure its quality and discoverability. It supports building interoperability across business, security, legal, semantic and technical themes.

Current state assessment establishes a strong understanding of your agency's business, information and data management environment.

Future state vision builds a vision of an improved future state and the requirements for creating new services and systems.

Gap analysis analyses and quantifies gaps between the current and improved future state. It forms the basis of requirements.

Planning and design plans how your revised business, information and data management environment will operate once the future state is implemented.

Implementation builds your agency's new interoperable state into its business using common tools and solutions. It also reminds you to monitor new processes for ongoing improvements.

Related links (naa.gov.au)

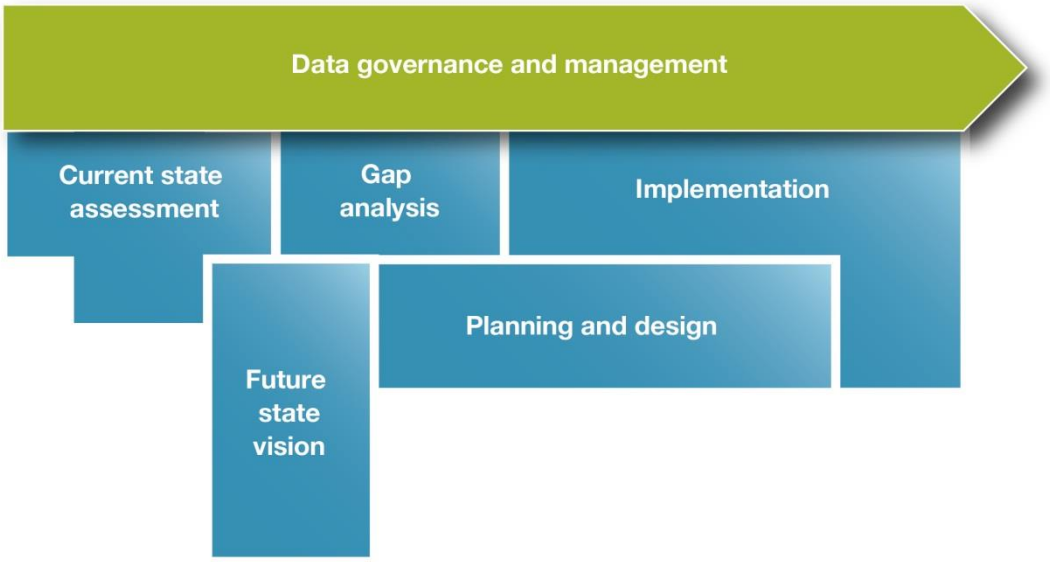
- Interoperability key themes
- Interoperability scenarios
- Interoperability scenarios pdf

4.1. Data governance and management

Data governance and management activities give you broad oversight of your data assets, along with managing them to achieve certain goals such as making day to day decisions to improve quality and discoverability. It is an essential component of information governance.

The key to good data governance and management is the consistent application of standards, policies and procedures. It supports building interoperability across business, security, legal, semantic and technical themes in your interoperability projects.

It should be applied at the beginning and throughout each of the interoperability development phases.



Some key activities include:

- Metadata for interoperability
- Data indexing and discovery
- Taxonomy
- Open data and formats
- Data quality
- Data compliance and security

Before commencing any work on building your interoperability, you need to become familiar with how your agency governs its information and data. An information and governance framework or similar can help you identify who is responsible for data governance and management standards for your agency.

Why data governance and management?

- ensure data has integrity and is available, usable, and secure
- conduct reviews and audits to identify what data needs to be interoperable based on their value and risk
- monitor information platforms, systems and applications for compatibility
- assess the risks of sharing (security, privacy, legal or other possible restrictions) and develop mitigation strategies or controls
- negotiate custody, ownership or shared access agreements
- negotiate the conditions on which information can be shared legally and ethically
- coordinate roles and responsibilities for stewardship, curation and maintenance of data.

Data disposal

Like all records, data is subject to legal requirements, standards and policies that determine its retention and disposal.

Data champion

A data champion is someone in your agency who is formally assigned rights and responsibilities to ensure data is being managed in line with relevant frameworks, policies and standards.

This may be a Chief Data Officer or a Chief Information Governance Officer who assures:

- data quality is fit for purpose
- the agency has accountable, authoritative, single source of truth data
- data sharing agreements are established with trusted users and agencies
- data access requests are streamlined
- data risks are managed
- data stewards are appointed to valuable data sets.

Related links (naa.gov.au)

- Chief information governance officer
- Compliant destruction of Australian Government records
- Disposing of information
- Information governance framework
- Information Management Standard
- Interoperability development phases
- Legislation, policies, standards and advice
- Managing information

Related information

- The ABS Data Quality Framework – Australian Bureau of Statistics (ABS)
- Data Management Association (DAMA)
- Information destruction and retention requirements – Australian Law Reform Commission (ALRC)
- Office of the National Data Commissioner (ONDC)

4.1.1. Metadata for interoperability

Metadata is structured information that defines and describes data. It plays an important role in ensuring users and systems understand the meaning of exchanged information and data.

Metadata exists in many sources within and external to your agency. Eg: documents, spreadsheets, databases, data models, legacy systems, social media and big data platforms.

You participate in creating metadata by using your agency's applications and tools, eg:

- electronic document and records management system (EDRMS)
- content management system (CMS)
- customer-relations management system (CRM)
- extract, transform and load tools (ETL)

Metadata standards ensure metadata is consistent, useful and understood over time. Establishing a minimum metadata set in your agency will also help you understand your user and business needs. This will help prioritise metadata for remediation and ensure that compliance with metadata standards is not an after-thought in developing new systems and processes.

Assessing metadata to see if it meets the standards for a specific process is a common task in building interoperability and may be included in your data quality assessment.

Metadata strategy

A metadata strategy assists in improving metadata governance across your agency. It documents current and future-state practices and how your agency manages its metadata.

Your strategy can also establish metadata standards that facilitate interoperability within your agency, between agencies and across jurisdictions. Eg, the metadata standard AS/NZS ISO 19115.1:2015 is used across jurisdictions to describe geographic information and services. This provides a consistent way to communicate information about the resources, objects and assets.

The metadata strategy links to your broader data and information governance environment, including your information governance framework.

Your agency should aim to manage and create metadata in a way that allows the metadata to be integrated. Ensuring metadata is relevant and kept up to date facilitates these processes and can be supported by your metadata strategy including the implementation of:

- roles and responsibilities for the creation and maintenance of metadata
- shared data dictionaries
- accurate tagging and identifiers
- change management processes.

Metadata harvesting

Metadata harvesting tools and protocols can assist agencies in indexing large batches of metadata records. Metadata harvesting uses automated tools to collect metadata descriptions from diverse sources such as catalogues, websites and other repositories. Open Archives Initiative protocol for Metadata Harvesting (OAI-PMH) is an example of a protocol or Application Programming Interface (API) for harvesting data. It supports aggregating data from multiple sources into one collection.

Metadata repository

A metadata repository is a data store for metadata. It is the aggregation of a wide range of metadata from across your agency.

Bringing your individual repositories together to develop a central repository of metadata can be beneficial. It allows consumers to look across all of your agency's available information from one point.

TIP: technologies such as ETL have their own metadata repository. Don't forget to integrate these when creating a central repository.

Metadata statements

Metadata statements are a detailed technical description of a dataset and can include information such as the:

- data custodian
- keywords (related subjects)
- unique identifier
- extent (the geographical area covered by the data)
- data quality
- limitations on how it is used.

Metadata statements can help users understand the dataset so they can make decisions to speed up and facilitate the use, publication and exchange of data.

Metadata for publication and exchange

Your agency can enhance metadata management to support publication and exchange of data by:

- implementing centralised metadata repositories
- updating metadata files to align with standards used at other agencies.

Many approaches to designing and implementing metadata exchange across your agency are available. Metadata architecture can be planned so that it facilitates this exchange. Consult with your internal specialists to determine which solutions work best for your agency.

Common architectural approaches include:

1. **Centralised metadata architecture**

   This involves copying metadata information from other applications and replicating it in a centralised repository. Users can perform global searches through a single application.

   Strengths:

- metadata information is accessed from one point
- opportunity for improving metadata quality by aggregating and transforming metadata sources into one standard
- prompt query retrieval
- manual metadata entry is possible.

   Limitations:

- complex maintenance and version control
- challenging tasks such as rapid replication or synchronisation of metadata
- custom code may be required to integrate metadata into the centralised repository's schema.

2. **Distributed metadata architecture**

   This consists of an application that retrieves data from source metadata in real-time, when a user requests information. There is no centralised repository. The intermediary application uses source catalogues to determine which repository to request information from.

   Strengths:

- no maintenance or version control required as the metadata is from the source
- processing is reduced as there is no metadata replication and queries are distributed among different sources.

   Limitations:

- metadata sources may not adhere to the same standards. Custom code may be required to retrieve the different metadata structures
- capture of additional metadata from external repositories can be difficult

3. Hybrid metadata architecture

This uses a combination of centralised and distributed metadata architecture. It provides both real-time access and allows manual entry of metadata to be added.

Strengths:

- metadata information is accessed from one point in real-time
- metadata can be added to the repository
- manual metadata entry is possible
- by adding metadata you can implement version control
- metadata quality can be improved by users.

Limitations:

- dependent on source metadata repositories being available.

Related links (naa.gov.au)

- The Australian Government Recordkeeping Metadata Standard (AGRkMS)
- Electronic document and records management system (EDRMS)
- Functions thesaurus
- Interoperability scenarios pdf
- Interoperability scenario: Data publication and exchange
- Interoperability: Data quality assessment
- Interoperability: ETL technologies
- Information governance framework
- Minimum metadata set
- What is metadata?
- YouTube: Meta… What? Metadata!

Related information

- Metadata standard AS/NZS ISO 19115.1:2015 – ANZLIC. The Spatial Information Council
- METeOR - Australian Institute of Health and Welfare (AIHW)
- Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)

4.1.2. Data indexing and discovery

Data indexing and discovery is about:

- quickly and efficiently finding data across an agency
- understanding the current data landscape of your agency
- making best use of existing data.

Efficient data indexing and discovery involves scanning repositories, databases and file stores for datasets. These tasks can be combined with other activities to help understand where existing data may reside and its file naming conventions.

Data indexing and discovery can be used to:

- speed up a data holdings audit or information review
- create and update a catalogue of datasets such as an information asset register.

It will assist you to:

- learn how data flows through an agency
- document the data lineage of datasets
- identify opportunities for improvements.

Data lineage documentation can include information on the business rules that are applied to the data as well as how the data flows or is transformed to be in its current state.

Through the use of data indexing and discovery tools, agencies can prioritise the maintenance and remediation of high value data.

Data profiling

Data discovery also includes profiling the data and assessing the data quality. Data profiling, visualisation and search and query tools can be used on large data sets to streamline assessments and provide insight into the root causes of issues with your data. These processes facilitate improvement of a system's architecture and your agency's procedures that detail how you procure, store, manage, use and dispose of data (DAMA, 2017).

Key data profiling tasks include:

- identifying differences between your data and what you assume it to be
- researching your current data flow and identifying areas for improvement
- profiling both source and target data. This helps you determine the data transformation required to match ideal standards of a specific initiative.

Catalogue your agency's data

Knowing what information and data you have in your agency is key to good governance. You should check your relevant governance framework for guidance on where to find the most up-to-date intelligence about your holdings.

You may have:

- an information asset register that identifies your assets, their potential value and possible risks
- a data catalogue that documents your agency's datasets and may include their systems, sources and locations
- a business systems register or software license register that may include details about systems' data – along with any associated system information management plans
- a metadata repository that stores your agency's metadata
- a risk register or system security plan that may also document data holdings.

Automation tools can be used to index your agency's databases, and can feed into existing registers or catalogues to streamline the process of cataloguing datasets.

Search tools

Other tools that allow users to search for data include:

- Application Programming Interfaces (API) and web services
- existing front-end applications
- indexing and searching commercial off-the-shelf solutions
- business intelligence (BI) tools
- intranet sites
- customer and staff portals for metadata repositories
- data catalogues that support machine-based querying.

Related links (naa.gov.au)

- Conducting an information review
- Information governance framework
- Interoperability: Data quality
- Interoperability scenario: Streamline business processes
- Interoperability: Web data services
- Metadata for interoperability
- Template for system information management plan

Related information

- DAMA-DMBOK: Data management body of knowledge, 2nd ed., 2017

4.1.3. Taxonomy

A taxonomy is a classification scheme for organising information and data into meaningful groups. It helps agencies organise their information into hierarchical relationships to ensure it can be efficiently searched, found and its meaning correctly interpreted.

Correct and consistent understanding of shared information i.e., semantics is a fundamental theme for enabling interoperability.

In its simplest form, the hierarchical structure of taxonomy is established using parent/child relationships between objects. It uses controlled vocabulary to name each object in relation to other objects.

Taxonomy uses hierarchy and controlled vocabulary to:

- manage synonyms
- ensure consistent application of business terms
- ensure business terms are correctly interpreted
- reduce information ambiguity.

Taxonomy creation

When creating a taxonomy consider using existing data definitions.

They facilitate:

- common terms and classes across agencies
- reducing necessary data transformations
- correct interpretation of data and information.

A functions thesaurus can be used in this manner. You can use thesauri that are available for Australian Government agencies to identify new and necessary controlled vocabularies when developing or updating your taxonomy.

Automation tools

Taxonomies can be created using automation tools. Automatic Taxonomy Construction (ATC) is a process that uses automated tools to generate taxonomy classifications from bodies of text.

ATC tools are helpful when you are:

- needing to create large ontologies
- flagging and correcting errors in existing taxonomies
- regularly generating large amounts of data that needs to be classified.

Ontology

Ontology describes an object in the same way as taxonomy - by its hierarchical position *however* it also describes an object by its relationships to other objects that are not in its linear hierarchy.

Eg, your family's pet cat could be represented using common language taxonomy as:

- carnivore
    - feline
        - domestic
            - Burmese.

An ontology may also associate your pet with concepts of popular YouTube content or Egyptian beliefs. When comparing taxonomy and ontology, ontologies are more in-depth, complex and represent richer relationships.

Languages

Ontology languages such as Resource Description Framework (RDF) and Web Ontology Language (OWL) are commonly used to construct ontologies that reflect the rich diversity and complex data of the Semantic web. They generate machine-readable language to encode knowledge on web pages, enabling their information to be queried and analysed. These ontologies help describe classes, attributes, relations, and events and can contain a collection of taxonomies. Example of OWL ontologies are showcased by the Australian Government Linked Data Working Group (AGLDWG).

Related links (naa.gov.au)

- Classifying information
- Functions Thesaurus
- Interoperability key theme: Semantic
- Interoperability scenarios pdf

Related information

- An introduction to the Semantic Web – Cambridge Semantics (United States)
- Linked Data showcase - Australian Government Linked Data Working Group (AGLDWG)

4.1.4. Open data and formats

To support interoperability your agency should consider making its data and data formats open.

Open file formats (open formats) are files that are:

- machine readable
- not 'locked' into a specific technology product or vendor
- freely accessed
- able to be reused with open source or free technologies.

Depending on you agency's technology, data may not be available in an open file format and may require conversion. Some technologies enable this natively or alternatively Extract, Transform, Load technologies (ETL) can be used to convert data into the desired format.

TIP: When using ETL to create, maintain or remediate metadata, use the opportunity to also convert data into an open format.

Examples of common open file formats include:

Types of data	Open file format
Databases	XML, CSV
Text	ODT, XML, JSON, HTML, RTF
Web Archive	WARC
Tabular Data	CSV, ODS
Geospatial Data	SHP, GeoTIFF, KML, WMS, WFS, WCS
Containers	GZIP, ZIP

Open data

Data is open when it is:

- published in a way that is easy to find
- freely available and accessible
- published with licenses that allow reuse.

If you make your data open you must consider legislation and policy to ensure you have met your privacy, security, ethical and access requirements.

You should consider open file formats when you are working on open data initiatives. The Government's Open Data Toolkit can help you in planning how to make your data open. The Digital Transformation Agency (DTA) provides guidance on using open data to improve services.

Standards bodies for open data include:

- International Organization for Standardization (ISO)
- Open Geospatial Consortium (OGC)
- Standards Australia

Open data maturity

Agencies can assess their open data maturity using a common model such as the NSW Government Open Data Model. This self-assessment tool is used to assess progress towards open data publication.

Related links (naa.gov.au)

- Interoperability: Extract, Transform, Load technologies (ETL)
- Interoperability scenarios pdf

Related information

- International Organization for Standardization (ISO)
- OGC standards - Open Geospatial Consortium
- Standards Australia
- Measuring open data maturity - NSW Government
- Open data - Department of Prime Minister and Cabinet (PM&C)
- Open data advice - DTA
- Open data toolkit - Australian Government
- The Open Data Handbook - Open Knowledge International

4.1.5. Data quality

Ensuring data quality should be seen as an ongoing improvement program that is managed throughout data's lifecycle.

Data Management Association (DAMA) defines common characteristics (dimensions) of data quality as:

- accuracy
- completeness
- consistency
- integrity
- reasonability
- timeliness
- uniqueness/deduplication
- validity.

Data quality management is a continuous process which involves managing data from its initial creation to its potential destruction. The quality of your agency's data should always be fit for purpose. You can support this by establishing a data quality strategy that facilitates proactive monitoring and managing of data quality. Eg, data quality assessments are embedded in data migration activities.

A data quality strategy should link to your broader data and information governance environment, including your information governance framework.

Data quality assessment

A good data quality strategy defines appropriate standards, requirements and specifications for data quality controls. This includes developing data dimensions relevant to your business needs to monitor, measure, and report on quality levels of your data.

Data quality assessment tells you how effective data is in meeting your stakeholders' requirements and also helps you prioritise remediation on high value datasets.

Data quality is assessed by measuring specific dimensions of your data.

They provide a:

- vocabulary for defining data requirements
- way to determine data quality assessment results
- metric for ongoing measurement and improvement (DAMA, 2017).

There are different dimensions that can be used to assess data quality, eg:

- common dimensions of data quality from DAMA's Body of Knowledge
- the Australian Bureau of Statistics (ABS) provides guidance on assessing against ABS dimensions, to determine the quality of *statistical* data
- ISO8000 - a global standard for data quality and enterprise master data. You can use this to inform your agency's data quality standards.

Data quality tools

Tools can be used as a guide to understand the different dimensions of data quality and generate data quality statements. An example is the NSW Government, data quality reporting tool that can be used to generate data quality statements in various document formats.

Tools that automate data profiling and cleansing are also available and can help your agency enhance large amounts of data.

These tools can:

- profile, clean and monitor data quality over time
- assist in the validation of data
- provide statistics on agencies data
- help to identify patterns and provide direction on future data remediation.

TIP: Data remediation can be achieved through the use of ETL software which can process data based on business rules and transform the data into the required format.

Poor data quality

Common culprits for poor data quality	Outcome
Incorrect data entry validation	Invalid data is entered into the database
Change in business rules	New rules are not correctly propagated throughout existing data
Changes to the source data structure	Third-parties implement changes without notifying downstream users; business rules are not updated on systems following notification of changes
Requirement for uniqueness of instances	Incorrect identifiers being created
Incorrect business rules being applied to data	Loss of data
Incorrect temporal information	Difficulty to identify latest version of information and data, resulting in duplication

Data quality and metadata

Good metadata is essential in understanding and assessing the quality of your data. Data quality assessments determine if your data meets the expectations of its consumers and metadata plays a key role in clarifying those expectations. Eg, you can look at a record's metadata to see if it meets format requirements or if it has been updated according to business rules.

Metadata can also be used to record data quality assessments; this means metadata repositories can be used for storing and sharing data quality assessment results across your organisation.

Your metadata and data quality teams can work closely together to develop these processes. Their combined expertise can ensure that business rules, measurements or issues related to data quality are documented, developed and managed as per your agency's data strategy.

Related links (naa.gov.au)

- Interoperability: Data migration
- Interoperability: Data profiling
- Interoperability: Data remediation
- Interoperability: Extract, Transform, Load technologies
- Interoperability scenarios pdf
- Information governance framework
- Metadata for interoperability

Related information

- Australian Bureau of Statistics (ABS) data quality statement checklist
- Data quality reporting tool – NSW Government
- DAMA – DMBOK: data management body of knowledge
- International Organization for Standardization (ISO)

4.1.6. Data compliance and security

Interoperability projects can only be realised when you have managed your information and data risks. All Government information and data is subject to legislation, policies and standards. Interoperability projects need specific attention to data compliance and security requirements relating to:

- data exchange mechanisms
- privacy and de-identification
- licensing for mixed, reused or derived datasets.

Data security is put in place to prevent unauthorised access to information. It is a fundamental theme for enabling interoperability and should be addressed as an enterprise wide initiative with an agency wide security strategy.

Data security requirements across Government which your agency must consider include the:

- Protective Security Policy Framework (PSPF) which includes requirements for sensitive and classified information
- Australian Government Information Security Manual (ISM) which is the standard that governs the security of Government ICT systems and includes information on access controls.

Secure data exchange

In addition to the PSPF, your agency can ensure your processes and systems meet criteria for secure data exchange by referring to the Digital Transformation Agency's (DTA):

- Trusted digital identity framework
- Gatekeeper Public Key Infrastructure Framework.

Data exchange security considerations include:

- access restrictions such as IP whitelisting, multi-factor authentication, security tokens and API Keys
- HTTPS secure connections
- encryption of data in transit and at rest
- tamperproofing data that is publicly exchanged
- strict password syntax checks and password resets
- encryption of all passwords
- data storage locations such as on premises and in the cloud
- security classifications.

Privacy and de-identification

Privacy and the de-identification of information must be considered when releasing information online. It is essential that all data released has undergone the necessary privacy and de-identification checks. The Office of Australian Information Commissioner (OAIC) provides the following information that can help you understand and meet these requirements:

- Australian Privacy Principles
- Privacy Principles fact sheet.

Licensing and terms of use

Compliance with licensing and terms of use of exchanged data is a legal requirement. To ensure you meet this requirement you must understand:

- what licencing and terms of use the data is under
- if the data uses other derived datasets
- necessary transfer of derived data's licence and terms of use, to new datasets.

Terms of use are applied to data to ensure users do not use the data out of context and for purposes other than intended. For example if a dataset does not reach a certain data quality there should be appropriate terms of use agreements in place to ensure it is not used out of context.

Related links (naa.gov.au)

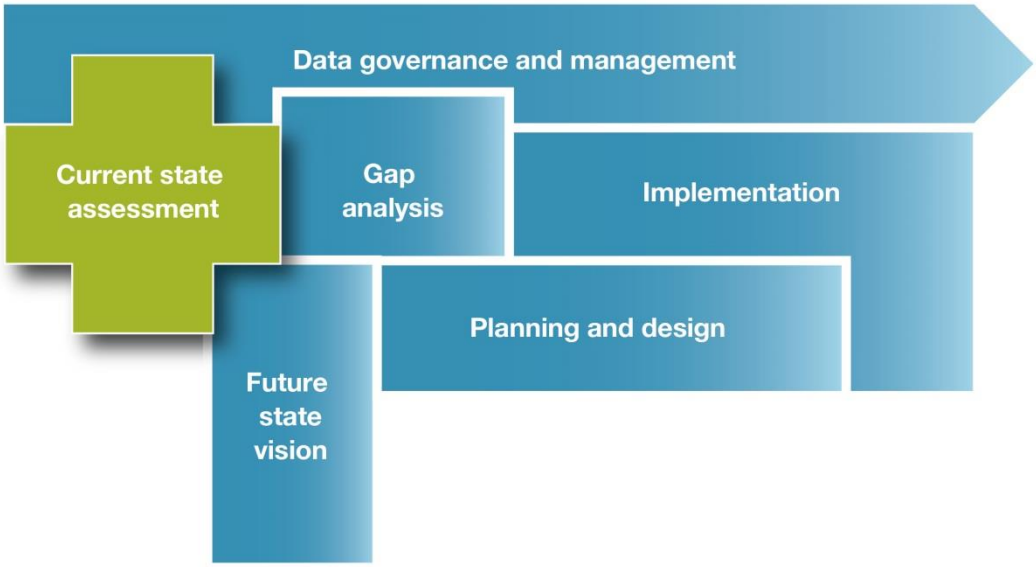
- Australian Privacy Principles and Commonwealth records
- Digital authorisations and workflows
- Interoperability key theme: Security
- Legislation, policies, standards and advice
- Protecting online information

Related information

- Australian Privacy Principles – OAIC
- Gatekeeper Public Key Infrastructure Framework - DTA
- Information Security Manual – Australian Signals Directorate (ASD)
- Privacy Principles fast sheet - OAIC
- Protective Security Policy Framework - Attorney-General's Department (AGD)
- Sensitive and classified information - AGD
- South Australian Information Security Management Framework
- Trusted digital identity framework - DTA
- Whitelisting - Wikipedia
- Multi-factor authentication - Wikipedia
- Security tokens - Wikipedia
- API Keys - Wikipedia

5.1. Current state assessment

Assessing your current state helps your agency to understand your business, data, and information management environment. This provides a strong start to determining your project's business needs.



When assessing your current state consider:

- the objectives and goals for the business processes and its data
- the current processes, procedures and practical methods of users – and what they want to achieve
- the hurdles, gaps and inefficiencies staff and users have encountered when using the business processes and managing data
- the legislation applicable to your agency
- results from activities that identify your systems, information and data eg, business systems assessment framework (BSAF) and data holdings audit.

Features and tasks

The approach to this phase differs across agencies. Your agency must determine the features and tasks that your current state assessment needs to build interoperability. Your agency's specific needs will define the method and priorities of your assessment.

Feature	Tasks
Data architecture and data flow	For the system being assessed, document: • the current technologies • the flow of data from collection points to general business use and storage • data structures, models and quality standards, and metadata architecture
Data/business system discovery	For general business: • refer to your existing systems register, data catalogue or information asset register • catalogue the existing business systems and data • use data harvesting software to find data across your agency's data collection • gather information on the legislation, policies and standards applicable to the data • gather information on the data's custodianship • undertake an audit of data holdings for data migration to determine the types and volume of data in your agency
Data profiling	Analyse and identify data that needs migration or remediation for a new solution. In a data analysis, include: • data format • data quality • data complexity • relationships with other datasets • collection points • storage and size • data lineage • security classification
Define business rules	Document the current business rules that apply to the data
Data quality assessment	Analyse the quality of data to assess whether it is fit for purpose.
Determine relevant standards and formats	Analyse the standards and formats your data must comply with. These include: • industry/sector standards for data exchange • data structure, formats and taxonomies • semantic web standards.

Key participants

Assessing your current state requires a range of skills from different sections across your agency. If your agency does not have the necessary time or internal expertise, external parties may be necessary to undertake the work for you.

Key participants may include:

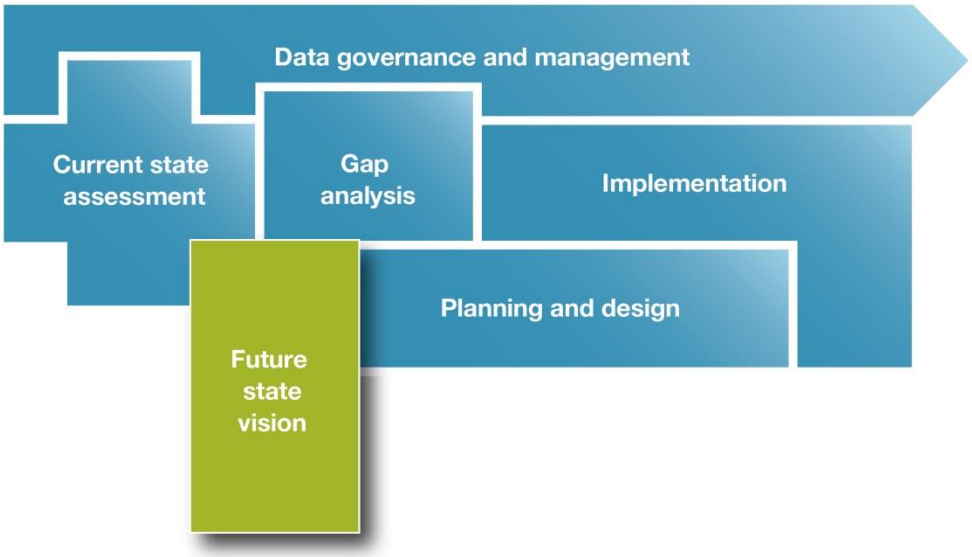
- project sponsor
- business and data analyst
- data / information architect
- data consumers
- security specialist
- data champion
- subject matter experts
- IT developer
- data specialist
- information governance manager
- chief information governance officer

Related links (naa.gov.au)

- Assessing information management functionality in business systems
- Chief information governance officer
- Interoperability: Data governance and management
- Interoperability: Data champion
- Interoperability development phases
- Interoperability scenarios pdf
- Legislation, policies, standards and advice

6.1. Future state vision

A future state vision helps your agency understand your requirements for creating improved services, systems or processes.



When developing the future state vision consider:

- researching user needs to engage directly with internal and external stakeholders and to understand how data may be used. Identify requirements such as data format, taxonomy or structure
- assessing the feasibility of the improved state against your agency's capabilities. A minimum viable product can help this process
- engaging consistently with the senior leadership team to gain the authority you need for decisions
- creating customer story maps (user stories) to understand a user's journey and their goals, actions, means and hurdles.

Key participants

The key participants required for a future state vision would be similar to those required for the current state assessment:

Key participants may include:

- project sponsor
- business and data analyst
- data / information architect
- data consumers
- security specialist
- subject matter experts
- IT developer

- data specialist
- information governance manager
- data champion
- chief information governance officer.

Related links (naa.gov.au)

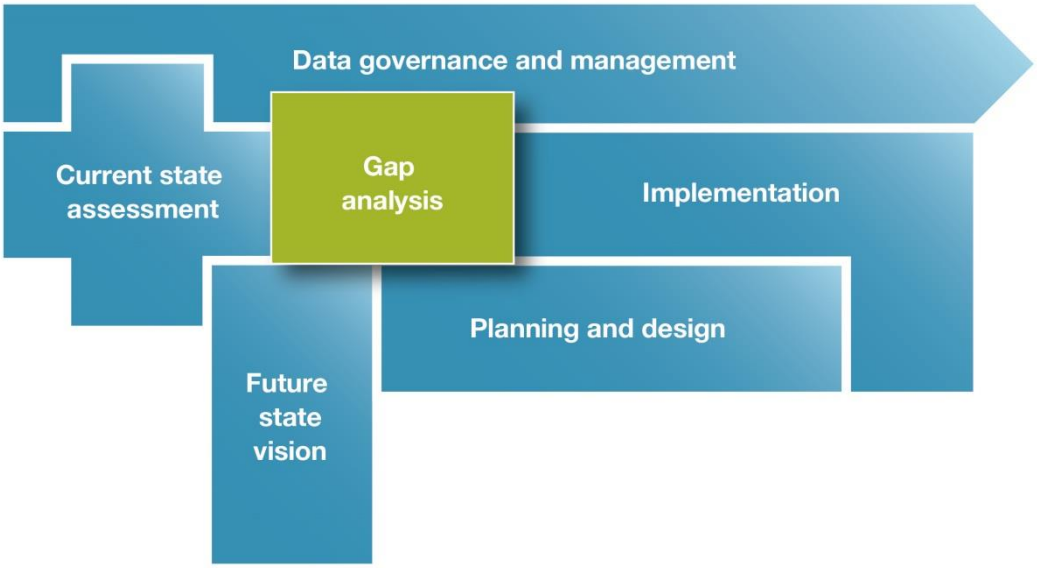
- Chief information governance officer
- Interoperability: Current state assessment
- Interoperability development phases
- Interoperability: Data champion
- Interoperability scenarios pdf

Related information

- Minimal viable product - Wikipedia
- User stories - Wikipedia

7.1. Gap analysis

The gap analysis phase gives you an opportunity to analyse and quantify gaps between the current and future state designs for improving data interoperability. A gap analysis needs to provide a holistic view at an architectural, application and/or data level. The results will form the basis of your research to find a solution. There may be issues with existing systems that you can only discover through detailed examination.



During a gap analysis you should consult end users to understand the pain points and issues they have with the current data and applications. For example, some data may not conform to existing business rules and may have some quality issues.

Once the gaps are identified, you can take appropriate action including remediation work as part of implementing the new application and migration. Before you implement new applications, your agency needs to approve the mapping specifications and supporting documents.

Key participants

The key participants required for a gap analysis would be similar to those required for the current state assessment:

Key participants may include:

- project sponsor
- business and data analyst
- data / information architect
- data consumers
- security specialist
- subject matter experts
- IT developer

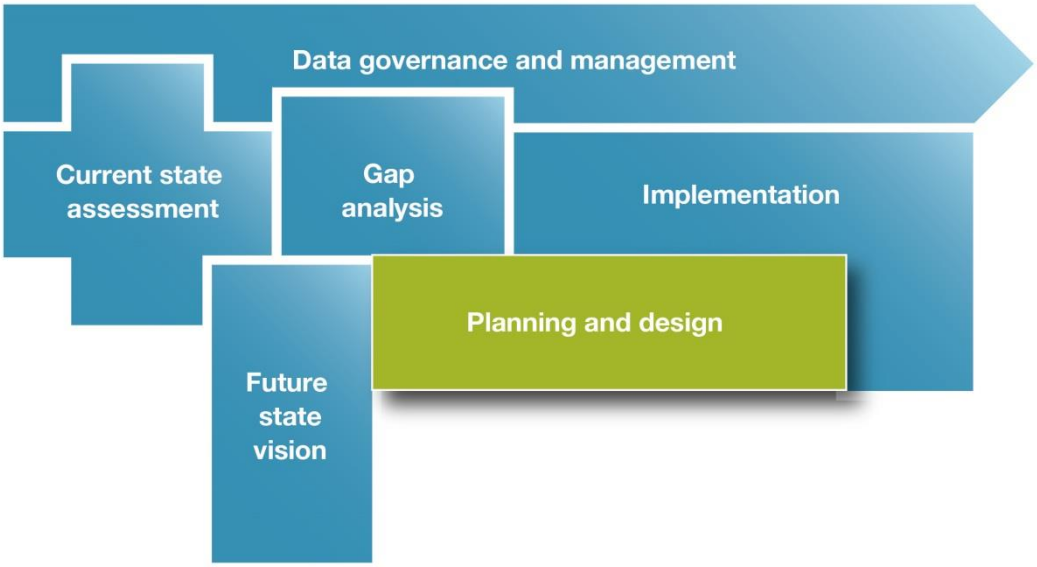
- data specialist
- information governance manager
- data champion
- chief information governance officer.

Related links (naa.gov.au)

- Chief information governance officer
- Interoperability development phases
- Interoperability: Current state assessment
- Interoperability: Future state vision
- Interoperability: Implementation
- Interoperability scenarios pdf

8.1. Planning and design

Planning and design is undertaken after you have analysed and quantified gaps between the current and future states. This phase plans how the new business environment will be structured and will operate in the future state vision.



Features and tasks

Agencies will approach this phase in different ways. Your project's gap analysis will inform your planning and design and address your specific requirements.

This phase can be challenging as you need to balance any new features and tasks with your agency's available resources and existing infrastructure.

The planning and design features and tasks are accomplished through researching user needs, consulting stakeholders and using internal skills and expertise. It is also the ideal stage to identify and assign roles.

Feature	Tasks
Data flow and data architecture	Document the required data flow for the system or agency and the business rules applied to the data. This includes the flow of data to and from: • collection points • data storage • third parties Document the solution that uses systems and architecture to: • prioritise using existing technologies or business processes before introducing new ones • adopt an enterprise-level approach by ensuring the number of different technologies in use across the agency is kept to a minimum.

	Deliver solutions that address: • storage, backup and recovery systems • scalability/availability • data retention periods • data structures and data models • metadata architecture.
Security designs	Document the state of access, authentication, security rules and processes required
Data integration plan	Document the way in which applications will communicate with each other and external parties. This includes: • interaction models. These detail how applications will interact with each other • data orchestration. These detail the solution's data flow.
Data migration plan	Document the process of moving data from the current to a new system. A successful data migration includes: • data profiling and analysis that is conducted in the current state assessment • data remediation work and data quality improvements as part of the necessary data transformation • migrated data that meets the standards for data in the new system • a planned approach to implementation. Decide if the data migration will be performed all at once, in stages, or if it will be continual as both legacy and new applications will be maintained.
System retirement plan	Document the decommissioning of legacy applications and data stores following the implementation of new applications for data integration or exchange. The system retirement plan must ensure that decommissioned applications and data meet storage and retention guidelines.
Metadata strategy	Document your agency's metadata management. This includes: • future state metadata architecture • data governance practices • data security • an implementation plan moving from the current to the future state design.

When designing an application or system your agency needs to consider:

- increasing data value through unified systems
- reducing data complexity
- making data more available
- ensuring data quality
- assessing data interoperability – whether it can be shared between agencies and with the public
- risks and dependencies.

Key participants

Planning and design requires a range of skills and competencies from different sections across your agency.

Key participants may include:

- project sponsor
- business and data analyst
- data / information architect
- data consumers
- security specialist
- subject matter experts
- IT developer
- data specialist
- information governance manager
- data champion
- chief information governance officer.

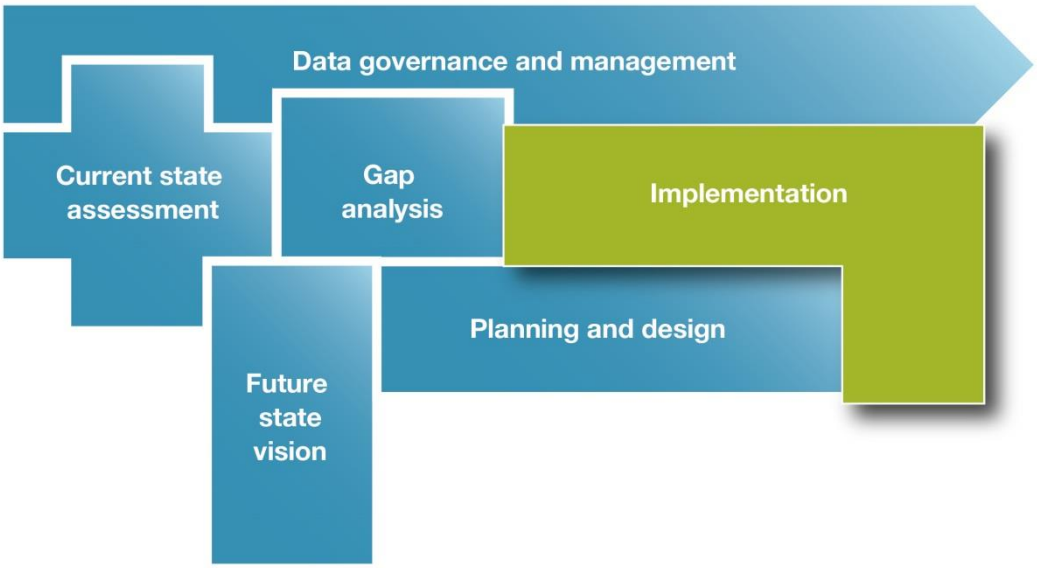
Related links (naa.gov.au)

- Chief information governance officer
- Interoperability development phases
- Interoperability: Current state assessment
- Interoperability: Data champion
- Interoperability: Future state vision
- Interoperability: Gap analysis
- Interoperability scenarios pdf

9.1. Implementation

This phase is about applying the right solutions and tools to develop interoperability. Implementation decisions will be informed by the planning and design phase that outlines how your revised business environment will operate.

Your implementation isn't complete until you have also considered how you will embed new processes into business as usual (BAU) – along with a plan for continuous improvement.



The solutions and tools summarised below are specific to the four interoperability scenarios that highlight data integration, legacy data migration, data exchange, and metadata for publication and exchange.

Data integration

Advice on approaches to combine and publish data from different sources into usable and trusted information.

Data collection

Advice on how data collection is used to gather information from users.

Data migration

Advice on transferring data from one application or format to another. Includes advice to address data quality and lineage issues during a data migration.

Data exchange

Outlines the importance of agreed data formats and structures between parties exchanging data. Includes advice on data exchange standards for the structure of data elements.

BAU, monitoring and improvement

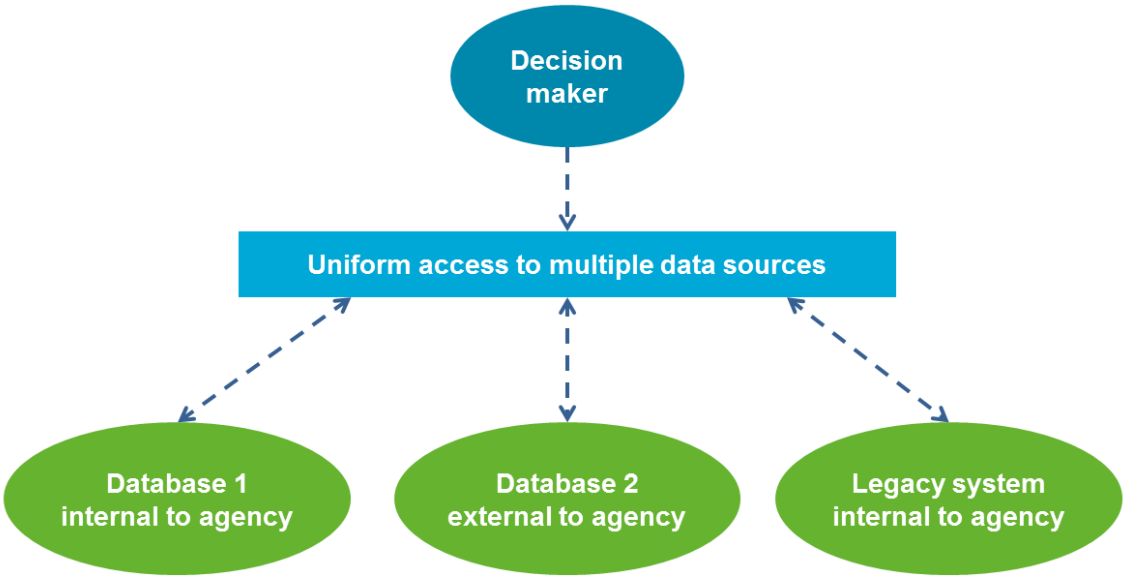
Advice to help you embed new processes into business as usual (BAU). This page also outlines how to monitor the solution for continuous improvement.

Related links (naa.gov.au)

- Interoperability: Planning and design
- Interoperability scenarios pdf

9.1.1. Data integration

Data integration is the process of combining data from different sources into usable and trusted information.



To engage technical experts in interoperability projects you need to understand technical aspects and language used in data integration, including:

- the various approaches to integrating data
- how web data services and APIs can help exploit data sharing
- various architecture models influence what tools and techniques you choose
- how Extract, Transform, Load (ETL) tools may help you.

There are several different ways of integrating data – whether it is within your agency, between agencies or with third parties.

Data integration approaches include:

Data integration approach	Description
Data Consolidation	Involves physically bringing data together from various sources to create one single data store. ETL is commonly used to consolidate data.
Data Propagation	Involves copying of data from one location to another and can often be a two-way data exchange between the source and target data store. Methods include: • Application Programming Interfaces (APIs) commonly used in a Microservices architecture • Enterprise Service Bus (ESB) used in a Service Oriented Architecture (SOA) to communicate data between the service provider and the service consumer • Enterprise Application Integration (EAI) • Cloud-based integration platforms such as Integration Platform as a Service (iPaaS) Large scale database data integrations often use the Enterprise Data Replication (EDR) method to copy and collect data from different source database systems and move that data to a broad range of target locations.
Data Virtualisation	Involves providing a single view of data by retrieving and interpreting data from different sources. This means that data can be viewed from one location but may be stored in other locations.
Data Federation	Involves using virtual databases to provide a single virtual view of data from various sources. Data federation provides users with a view to a collection of data sources regardless of their structure. Enterprise Information Integration (EII) is a technology that supports data federation.

Web data services

Web data services or web services helps improve interoperability by enabling machine-to-machine interaction, allowing data to be shared and re-used over a network (web) internal or external to your agency. Interoperability is achieved through XML-based standards that provide a common way to define, publish, and use web services.

Your agency may use web data services with other agencies or third parties eg, to make performance reporting processes more efficient. Their data can be transmitted to your agency using system to system transfers that comply with a web service exchange specification.

TIPS: when you design or implement new data services, you should always check there are no existing services that can be used or reused to meet the service's requirements.

Data integration should be based on standards so that services can easily be interchanged by your agency without the end user having to make any amendments on their systems.

Application Programming Interfaces (APIs)

An API is a set of rules and specifications for a software program to follow to facilitate communication across applications. Similar to web services, an API's end goal is to facilitate the integration and sharing of information and data. APIs are often referred to as the broker between two applications that controls what data can be shared through requests by other parties. Representational State Transfer (RESTful) or Simple Object Access Protocol (SOAP) are common APIs**.**

RESTful APIs are increasingly used in government settings and are also an efficient way for you to package your data for the public and third parties to unpack and consume.

APIs can be used by your agency to:

- streamline business processes across business lines, agencies and third parties e.g. the Australian Tax Office uses APIs with third parties to streamline the collection of data for tax related activities
- share your information and services to the public and third parties e.g. APIs allow the Australian Bureau of Statistics to transfer statistical data from the ABS to user's machines or applications
- help collaborate with other agencies to build solutions on top of your current systems.

API management

API management is responsible for handling user access such as API access keys or tokens. Security provisions can be built into an API management to control data access.

TIP: when designing APIs, consider their scalability, availability and reuse potential within and external to your agency. It is also important to ensure they are well defined with information and guidance made available on the options and parameters to use the APIs.

Useful guidance on developing APIs includes:

- Application Programming Interfaces - Digital Transformation Agency (DTA)
- Application Programming Interfaces for software developers - Australian Securities & Investments Commission (ASIC)
- Application Programming Interfaces – Government (United Kingdom)

Architecture models

Architecture models are the building blocks and influence the way you design your web data services and APIs. Popular architecture models include:

- **Service oriented architecture (SOA)** is a software architecture where distinct components of the application provide services to other components via a communications protocol over a network. SOA have historically been used in larger, complex enterprise environments that require many different types of integrations.
- **Microservices architecture** is a software architecture that is made up of small, independent processes which communicate with each other using APIs. Each service is created to serve only one specific business function and is completely independent of other services. It is commonly used in smaller application environments that integrate with other APIs such as developing a mobile or web application.

Extract, Transform, Load (ETL) technologies

ETL is traditionally used for moving data within and across applications and agencies. ETL tools can be used in the creation of web data services to **create workflows** to transform and deliver the data in the required format, taxonomy and structure.

ETL technologies can be used to:

- collect data from various sources with potentially different taxonomies, formats and structures
- check for new updates to data and process it to an agency's master database - while performing automated validation checks on the data to ensure it have been received in the correct format and quality
- transform the data according to business rules eg, into structured and validated datasets
- load the data into the target destination database or application.

TIP: ETL may be a slower integration method – consider other methods that suit your integration project.

Related links (naa.gov.au)

- Interoperability scenario: Streamline business processes
- Interoperability scenarios pdf

Related information

- Application Programming Interfaces – DTA
- Application Programming Interfaces for software developers – ASIC
- Application Programming Interfaces – Government (United Kingdom)
- Cloud-based integration - Wikipedia
- Enterprise Information Integration – Wikipedia
- Enterprise Service Bus - Wikipedia
- Replication - Wikipedia
- Representational State Transfer (RESTful) - Wikipedia
- Simple Object Access Protocol (SOAP) - Wikipedia

9.1.2. Data collection

Data collection is the process of gathering information from users. Your agency should obtain information from internal and external users in the required digital form. This will allow the data to be stored directly in the database without manual intervention as the data should be in the required structure.

Data collection methods include:

- Application Programming Interfaces (APIs)
- mobile applications
- online customer portals
- File Transfer Protocol (FTP) and Secure File Transfer Protocol (SFTP)

Changing collection methods

During the gap analysis or planning and design phases, you may identify the need to improve data collection methods to satisfy your future vision.

Following the implementation of a new data collection method, maintaining both the old and new data collection methods can be beneficial for a short period of time to ensure that the new systems integrate correctly and the data is of the desired quality. When your agency is confident with the new data collection method, the legacy collection mechanisms can be decommissioned.

Related links (naa.gov.au)

- Interoperability: Application Programming Interfaces (APIs)
- Interoperability: Future state vision
- Interoperability: Gap analysis
- Interoperability: Planning and design
- Interoperability scenarios pdf

Related information

- File Transfer Protocol (FTP) - Wikipedia
- Secure File Transfer Protocol (SFTP) - Wikipedia

9.1.3. Data migration

Data migration is the process of transferring data from one application or format to another. It is often required with the implementation of a new application, which may require data to be moved from an incompatible proprietary data format to a format that is future proof and can be integrated with new applications.

Data migration considerations and activities

Understand your data and its quality

A data migration can identify if data from an old application is not up to the quality that is suitable for a new application. Before you start a migration activity, you should refer to your data catalogue or information asset register. A recent data holdings audit or information review may also be required to understand the data to be migrated. This can be assisted by automated indexing and discovery tools.

TIP: it is critical to understand where the master data or authoritative source is held prior to the data migration to ensure the latest version of the data is being migrated.

Following a data holdings audit a data quality assessment can be undertaken to understand the quality of data and determine the rules and actions required. Data profiling can help you determine the quality of data, eg relevance, format, consistency, validity, complexity, completeness, accuracy, accessibility, compliance and structure of the data. Automated data profiling tools can be used to streamline this process, especially when there is a large quantity of data to profile.

Identify stakeholders

Prior to starting the transfer of data it is important to ensure the necessary planning, profiling, and migration plans have been created and approved by the business. Data migration can only be a success if business is engaged throughout the migration, and requires business stakeholders ranging from the senior leadership team to data analysts.

You can start by identifying team members that need to be involved in the migration, testing, auditing, review and sign-off stages.

TIP: ensure that end users are also involved in business rule validation and testing throughout the migration process.

Data extraction

This stage involves copying or moving the data from legacy stores to a secure location to have the data prepared for migration. During this process data aggregation may be required to bring together several datasets to make the data more meaningful and fit for purpose for the new application. This may involve the creation of a new combined master dataset which maps out the data linkages between the different datasets to be aggregated.

TIP: ensure the necessary backups have been put in place in the event of data corruption.

Data remediation – improving the quality of data

Data migration processes can provide an opportunity to improve the quality of the data. Once the quality of the data has been measured, certain actions can be applied to remediate data. Data remediation can occur before, during or after the transfer of data to the new application. Extract, Transform, Load (ETL) tools are often used to categorise and improve data quality.

An example of data quality actions before, during or after a data transfer:

Data quality action	Description
No Action	Data issues are small and not meaningful and will not cause a problem post migration.
Remediate during the ETL processes	Data issues should be remediated during the ETL transformation of data from the legacy to the new application
Remediate in source databases	Data issues should be remediated in the source database
Remediate in target application	Data issues should be remediated once the data has been loaded into the application. This may cause some complications as the data may not pass target validation or have errors using the new business rules.

Data migration techniques

A wide range of techniques can be used to perform a data migration. The level of automation that is possible during a data migration will depend on the maturity of your data. For example it may be necessary to first scan physical documents for capture before Optical Character Recognition (OCR) software can be used to convert scanned copies to digital structured data.

Once data is in a digital format potential data migration technologies could involve:

- Optical Character Recognition (OCR) software for character recognition and digitisation
- Extract Transform Load (ETL) software for data transformation such as format conversions
- OGR2OGR software for spatial data migrations
- Machine Learning (ML) software for automated mapping of data structures for migration.

Data can be transferred in one go, or in stages to ensure quality is maintained and can be completed in smaller agile sprints.

Data validation, auditing and verification

Testing and quality check points should be occurring throughout the data migration with all results recorded for auditing. Unit, application, system and volume tests should be undertaken as early as possible in the migration to ensure there is time to update any code or business rules.

TIP: all new processes and data transformations should be documented to ensure there is traceability and auditability of the data. The migrated data needs to demonstrate the data is still authentic so it can be relied on as evidence.

Automated validation should also be considered to check the volume and quality of the data being migrated to ensure no data is lost and the data is fit for purpose. Ensure any subsequent metadata, lineage and data quality statements are updated with the data migration.

TIP: consider running legacy and new systems concurrently until you are confident with the new process and data. When stakeholders are confident with the new application the legacy system can be decommissioned.

Related links (naa.gov.au)

- Conducting an information review
- Interoperability: Data indexing and discovery
- Interoperability: Data quality assessment
- Interoperability: Extract, Transform, Load technologies
- Interoperability scenarios pdf

Related information

- OGR2OGR - Geospatial Data Abstraction Library (GDAL)
- Optical Character Recognition (OCR) - Wikipedia
- Machine Learning (ML) - Wikipedia

9.1.4. Data exchange

When you move data from one application to another, the way the data is structured may not be compatible between the two systems. Data exchange is the transformation or restructure of the data to suit the new system. The aim of a data exchange is not to lose or change the meaning of the information in the transfer.

Exchange standards

Data exchange standards are formal rules for the structure of data elements. A data exchange specification is a common model used by agencies and standardises the format in which data will be shared. The International Organization for Standardization (ISO) develop data exchange standards such as ISO 20614:2017 Data exchange protocol for interoperability and preservation.

Exchange agreements

Your agency should develop an agreed exchange format and structure with agencies you are supplying and receiving data from. This reduces time and effort required in restructuring or reformatting data. Agreed data exchange formats and structure also ensures users have a better understanding of the data and can reduce manual requests.

An example of an agreed data exchange is the National Information Exchange Model (NIEM) which is a common vocabulary that enables efficient information exchange across diverse public and private organisations. NIEM ensures that information carries the same consistent meaning across various communities by using a dictionary of agree terms, definitions, relationships and formats regardless of how the information it stored.

Data exchange languages

Data exchange languages facilitate how applications interpret each other's data. Common data exchange languages include RDF, XML, JSON, YAML, REBOL and Gellish.

Platforms for data exchange

When looking to implement data exchange solutions, agencies should first look to try and reuse existing government platforms and data hubs. For example data.gov.au provides government agencies an easy way to find, access, publish, and reuse public data. The Digital Transformation Agency (DTA) is also working on establishing common platforms for government to help your agency work with other agencies to deliver joined-up services for citizens.

If an existing solution is not available agencies should look to conduct a technology evaluation to find the most suitable technology. A popular mechanism of data exchange is through the use of Application Programming Interfaces (API) or web data services.

Related links (naa.gov.au)

- Interoperability: Application Programming Interfaces
- Interoperability: Open data and formats
- Interoperability: Web data services
- Interoperability scenarios pdf

Related information

- Data.gov.au – Australian Government
- ISO 20614:2017 Data exchange protocol for interoperability and preservation - ISO
- National Information Exchange Model - (NIEM; United States)
- Open data advice - DTA
- Platforms for Government - DTA
- Data exchange languages - Wikipedia

9.1.5. Business As Usual (BAU), monitoring and improvement

When building interoperability, new processes need to be documented and provided to users in an easy to access format, eg, through team collaboration software.

To ensure the maintenance of the new processes it is also important to:

- ensure responsibilities are well defined and documented so everyone is clear on their role
- undertake quality checks for potential user issues and barriers – which can also inform future improvements.

Monitoring and continuous improvement

During the implementation of data interoperability solutions and tools consider:

- integrating real-time data monitoring and analytics within the solution that will enable data to be collected on potential issues and accomplishments – even after the solution has passed testing and gone live
- analysing data from the solution to understand trends in data exchanges, data quality and quantity as well as monitoring business rule violations. This information can then be used to see if the solution is performing as expected and identify areas for improvements.

Some features to monitor include:

- how many users are using services and during what times of the day
- where are users originating from and using which technologies
- how long users are spending on the services and the success rate of them finding what they are after

- which services are being used the most
- which terms are being searched and generate poor results
- quality of the data
- logs from the applications to ensure there aren't any errors or issues
- volume of the data being exchanged and the speed at which data is exchanged.

Related links

- Interoperability: Data quality
- Interoperability: Data exchange
- Interoperability scenarios pdf
- Interoperability: Search tools