



Metadata Documentation Policy

1. BACKGROUND

In the course of its operations, WA Health collects, stores, uses and discloses a large volume of data. This is an important resource used for determining the clinical care of patients and for making decisions in relation to the funding, management, planning, monitoring, research and evaluation of health and health services in the State.

To manage the vast amount of data contained within WA Health Enterprise Systems, WA Health must maintain adequate documentation about these systems. Metadata is a documentation methodology that provides users with information about the purpose, processes and methods involved in the data collection.

This document provides the policy, processes and governance necessary for building and managing a standardised and centralised corporate repository of metadata.

2. SCOPE

This policy applies to the data contained in WA Health's Enterprise Systems only.

Enterprise Systems are large-scale, integrated information systems that support processes, information flows, reporting and data analytics across WA Health. Typically Enterprise Systems are classed as tier 1 applications requiring 24 hour, 7 day per week availability and technical support.

Enterprise Systems for WA Health include: Advanced Incident Management System; Alesco Human Resource Management System; Alesco Payroll System; Allied Health System ; Emergency Department Information System; Health Care and Related Information Systems – Client Management System; iPharmacy; Lattice Payroll System; Lattice Resource Management System; Oracle Financial and Supply Systems; Theatre Management System; The Open Patient Administration System (TOPAS); Stork and webPAS.

For a full listing of Enterprise Systems refer to the [WA Health Information Register](#).

3. PURPOSE

To manage the volume of data within WA Health Enterprise Systems, a metadata repository will be developed. The repository will summarise and store information about data that is collected within WA Health's Enterprise Systems. This is to establish a common understanding of the meaning of the data, to ensure its correct interpretation and use.

Metadata standards will be developed and ultimately adopted by all Enterprise Systems to ensure consistent definitions are applied for business, operational and legislative requirements.

A metadata repository will enable WA Health to:

- support the long term management of WA Health's data
- make data accessible, fit for purpose and sustainable
- gain a better understanding of the nature and types of data collected.

Due to the complexity and diversity of data collections within WA Health, a staged approach to developing the metadata repository shall be adopted. The four stages to building a WA Health metadata repository are as follows:

1. Document information about data within WA Health's Enterprise Systems.
2. Review WA Health's metadata documentation for completeness.
3. Develop WA Health's metadata model, including standardisation of common metadata items.
4. Develop governance, training and communication strategies for WA Health's data collections.

The purpose of this policy is to outline Stage One: 'Document information about data within WA Health's Enterprise Systems'.

4. POLICY

4.1 STAGE ONE - DOCUMENT INFORMATION ABOUT DATA WITHIN WA HEALTH'S ENTERPRISE SYSTEMS

Knowledge of the current data within each Enterprise System will assist in the development of a WA Health corporate metadata repository. Documenting existing metadata will also allow the development of standardised items and provide consistency to information produced by WA Health.

Existing Enterprise Systems

Data Custodians must document the information contained within each Enterprise System. A template (Appendix A) has been developed to facilitate this documentation and will ensure all pertinent information is captured. The template will be available in Excel format to assist in completion.

The template titled 'Data Dictionary' contains the basic definitions in relation to data collection. The definitions assist in understanding the meaning of data to ensure the correct use by its owners and users.

The data dictionary template is in three parts:

- **Data set specifications** - Specifies basic information about the Enterprise System e.g. TOPAS. The data set specification defines the characteristics of the collection e.g. the scope of TOPAS.
- **Value domains** - A description of each of the items within the data collection, to be used for reporting and analysis e.g. triage, departure destination, major diagnostic category.
- **Quality declaration** – Provides information on a range of aspects of the quality of the data collection. The quality declaration assists in assessing the quality of the

[Quality Policy \(OD 0380/12\)](#) will assist in completing this quality declaration.

When completed, a copy of the data dictionary must be sent to the Information Development and Management Branch within the Performance Activity and Quality Division (PAQ), allowing Stage Two, a review of WA Health's current metadata documentation, to commence.

New Enterprise Systems

As part of the development stage for a new Enterprise System, the Data Custodian must complete a data dictionary. The data dictionary will assist in the data collection approval process by enabling a greater understanding of the content and purpose.

The completed data dictionary must be:

- Submitted with the 'Assignment of Data Custodian' template within the [Data Stewardship and Custodianship Policy \(OD 0321/11\)](#) to the Enterprise System Data Steward for endorsement.
- Sent to the Information Development and Management Branch, PAQ, for inclusion in Stage 2: 'Review WA Health's current metadata documentation'.

5. DEFINITIONS

Data Custodian – The person(s) responsible for the day-to-day management of data from a business perspective. The Data Custodian aims to improve the accuracy, usability and accessibility of data within the data collection.

Data Dictionary – Documents the basic definitions of all aspects of the data collection. It contains the list of value domains that are available in the data collection and the information about the value domains. A data dictionary will assist in understanding what data is available and what it represents.

Data Stewards – A delegated person responsible for setting the overall strategic direction of a specific data collection. They ensure the collection is developed, maintained and utilised in accordance with the strategic goals of WA Health. Data Stewards are also responsible for authorising access, use and disclosure of data from the collection for clearly defined purposes that comply with WA Health's statutory obligations.

Enterprise Systems - Large-scale, integrated information systems which support processes, information flows, reporting and data analysis across WA Health. Typically enterprise systems are classed as tier 1 applications requiring 24 hour, 7 day per week availability and technical support.

Metadata Repository - Stores and catalogues metadata. The metadata within the repository will reflect the standard definition for all WA Health data, which will ensure consistent reporting definitions for business, operational and legislative requirements.

6. ASSOCIATED POLICIES

[Data Collection Policy \(OD 0381/12\)](#)

[Data Quality Policy \(OD 0380/12\)](#)

[Data Stewardship and Custodianship Policy \(OD 0321/11\)](#)

[Information Access and Disclosure Policy \(OD 0360/12\)](#)

[Information Lifecycle Management Policy \(OD 0371/12\)](#)

[Information Use Policy \(OD 0390/12\)](#)

7. RELEVANT LEGISLATION

Health Act 1911

Health Services (conciliation and review) Act 1995

Hospital and Health Services Act 1927

Human Reproductive Technology Act 1991

Mental Health Act 1996 (WA)

Freedom of Information Act 1992

Financial Management Act 2006

Public Sector Management Act 1994

State Records Act 2000

8. SUPPORTING DOCUMENTS

Australian Institute of Health and Welfare. METeOR. About Metadata.

<http://meteor.aihw.gov.au/content/index.phtml/itemId/268284>

Bargmeyer, B.E. and Gillman, D.W. Metadata Standards and Metadata Registries: An Overview. <http://www.bls.gov/ore/pdf/st000010.pdf>

AS ISO/IEC 11179.1. 2005. Information Technology – Metadata Registries (MDR) – Part 1: Framework.

Data Dictionary

Below is an example of a completed data dictionary template. An electronic excel version will be available to assist completion.

Data Set Specification Template - Specifies a group of value domains (data items) and the conditions under which this group is collected e.g. The Open Patient Administration System (TOPAS). The Data Set Specification defines the characteristics of the collection e.g. the scope of TOPAS.

Data Set Specification											
Identifying and definitional attributes											
Name	A single or multi-word designation assigned to a data set specification e.g. The Open Patient Administration System (TOPAS).										
Scope	A description of the circumstances under which the collection of specified data are required or recommended e.g. TOPAS contains patient and client episode information and activity. It is used in the metropolitan area and Bunbury Hospital.										
Collection and usage attributes											
Guide to use	Any comments, advice or instructions for the interpretation or application of the Data Set Specification.										
Collection methods	Comments, advice or instruction for the actual capture of data e.g. TOPAS information is created at points of care throughout the hospital, including emergency, admissions and outpatients. Information relating to a patient's episode of care and demographic details are entered into the TOPAS system by hospital staff.										
Implementation start date	<table border="1"> <tr> <td>0</td><td>1</td><td>/</td><td>1</td><td>1</td><td>/</td><td>1</td><td>9</td><td>9</td><td>7</td> </tr> </table> The date upon which the collection of data for this specific version of the data set specification was first implemented	0	1	/	1	1	/	1	9	9	7
0	1	/	1	1	/	1	9	9	7		
Implementation end date	<table border="1"> <tr> <td></td><td></td><td>/</td><td></td><td></td><td>/</td><td></td><td></td><td></td><td></td> </tr> </table> OR, if ongoing please cross box <input checked="" type="checkbox"/>			/			/				
		/			/						
Comments	Any additional information that adds to the understanding of the data set specification.										
Source and reference attributes											
Source	Source(s) of the data within the collection and/or is the data a by-product of an existing system(s) e.g. TOPAS data is the source data and is not reliant on information from other systems.										
Submitting establishment/section/unit	The establishment/section/unit within WA Health responsible for the data set specification.										
Data Custodian	Data Custodian name for the submitting establishment/section/unit e.g. TOPAS, Head of Department, Patient Information Management Service, Sir Charles Gairdner Hospital.										
Data Steward	Name of Data Steward for the submitting establishment/section/unit e.g. TOPAS - Executive Director, PAQ.										
Governance attributes											
Purpose of the data	The reason for collecting the data e.g. TOPAS data is used for State and national reporting.										
Principal user(s)	The Principal users of the information e.g. Minister of Health, Community, Australian Institute of Health and Welfare and WA Health data collections.										

Value Domain Template

A description of the data items within the data collection used for reporting and analysis e.g. Triage category, departure destination, major diagnostic category. A Value Domain Template must be completed for each data item within the collection.

Value Domain													
Identifying and definitional attributes													
Identifier	(office use only) <input type="text"/>												
Name	A single or multi-word designation assigned to a metadata item e.g. Triage category.												
Synonymous name(s)	One or more synonyms for the metadata item name within the context of the given item e.g. Triage Severity Code.												
Definition	A concise statement that expresses the essential nature of the value domain and how it is differentiated from other metadata items e.g. The urgency of the patient's need for medical and nursing care, as represented by a code.												
Context	A designation and/or description of the application environment or discipline in which the value domain has meaning e.g. Non-admitted patient emergency department care.												
Classification scheme	The name of the classification scheme which is implemented for this value domain e.g. Not applicable for TOPAS triage category.												
Representation attributes													
Representation class	The classification of representation e.g. Code.												
Data type	A set of distinct values characterised by properties of those values and by operations of those values e.g. Number.												
Format	The presentation of values, including specification and layout of permitted characters and precision e.g. N.												
Maximum char length	The maximum number of characters permitted to represent the values e.g.1.												
Permissible values	<p>A list of codes plus code descriptions representing values on the primary collection form e.g.</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Meaning</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>Resuscitation: immediate (within seconds)</td> </tr> <tr> <td>2</td> <td>Emergency: within 10 minutes</td> </tr> <tr> <td>3</td> <td>Urgent: within 30 minutes</td> </tr> <tr> <td>4</td> <td>Semi-urgent: within 60 minutes</td> </tr> <tr> <td>5</td> <td>Non-urgent: within 120 minutes.</td> </tr> </tbody> </table>	Value	Meaning	1	Resuscitation: immediate (within seconds)	2	Emergency: within 10 minutes	3	Urgent: within 30 minutes	4	Semi-urgent: within 60 minutes	5	Non-urgent: within 120 minutes.
Value	Meaning												
1	Resuscitation: immediate (within seconds)												
2	Emergency: within 10 minutes												
3	Urgent: within 30 minutes												
4	Semi-urgent: within 60 minutes												
5	Non-urgent: within 120 minutes.												
Supplementary values	A list of codes and code descriptions representing values produced in the data cleansing process e.g. TOPAS has no supplementary values.												
Unit of measure	State the unit of measure e.g. Time.												
Collection and usage attributes													
Guide to use	Comments, advice or instructions for the interpretation or application of the value domain e.g. Permissible values are mandatory.												

Quality Declaration

The Quality Declaration assists in assessing the quality of the data for reporting purposes. The Data Quality Assessment Plan within the [Data Quality Policy \(OD 0380/12\)](#) will assist in completing this Quality Declaration.



Government of Western Australia
Department of Health

Quality Declaration

The Quality Declaration is intended to be an informative statement about the quality of the data set specification using seven dimensions of the data quality framework. Quality Declarations will complement, but not replace, any existing explanatory and technical notes, and more detailed documentation.

To assist in completing the Quality Declaration, refer to the Data Quality Policy 'Assessment Plan' (OD 0380/12). Completion of the Assessment Plan will assist in completing the below Quality Declaration.

When completing this section, please provide a brief summary, as well as the rating from the Assessment Plan.

Relevance

Relevance reflects the degree to which a data collection meets the current and potential needs of users. The purpose of the relevance dimension is to assess how well a data collection can adapt to change and whether the collection is perceived to be valuable. Refer to Criteria 38 to 42 within the Assessment Plan.

Timeliness

Timeliness refers primarily to how up to date the data is at the time of release. Timeliness is closely associated with relevance, in that if the data is not timely, the data may no longer be relevant for the needs of users. Refer to Criteria 21 to 23 within the Assessment Plan.

Accuracy

Accuracy refers to how well information captured within an information system reflects the reality. Refer to Criteria 1 to 20 within the Assessment Plan.

Accessibility

Accessibility refers to the ease with which the information can be obtained. Refer to Criteria 32 and 33 within the Assessment Plan.

Interpretability

Interpretability is similar to Usability within the Assessment Plan. Usability reflects the ease with which data from a data collection may be understood and accessed. Refer to Criteria 32 to 37 within the Assessment Plan.

Coherence

Coherence is similar to Comparability within the Assessment Plan. Comparability is defined as the extent to which the data collection is consistent over time and uses standard conventions, making it similar to other relevant data collections within the Department. Refer to Criteria 24 to 31 within the Assessment Plan.