

Appendix 1: Metadata standards

The *National community services data dictionary* is based on the 2003 version of the ISO/IEC 11179 standard for metadata registries. In this standard, a data element is specified in a number of subcomponents. These are:

- data element concept
- value domain
- object class
- property.

METeOR and the dictionary also use ‘Classification scheme’ which is another subcomponent of the standard, but it is used in a slightly different way from that envisaged by the standard.

In METeOR and the dictionary, these types of metadata (data that describe other data) are applied to the specification of data standards. It was found that an extension (an addition that does not contravene the existing standard) to the standard was required to accommodate other types of metadata used in the dictionary. This extension was required for ‘glossary items’ and ‘data set specifications’.

The structure underlying a data element in the dictionary is illustrated in Figure 1 (differences from the ISO/IEC standard are shown with dashed lines).

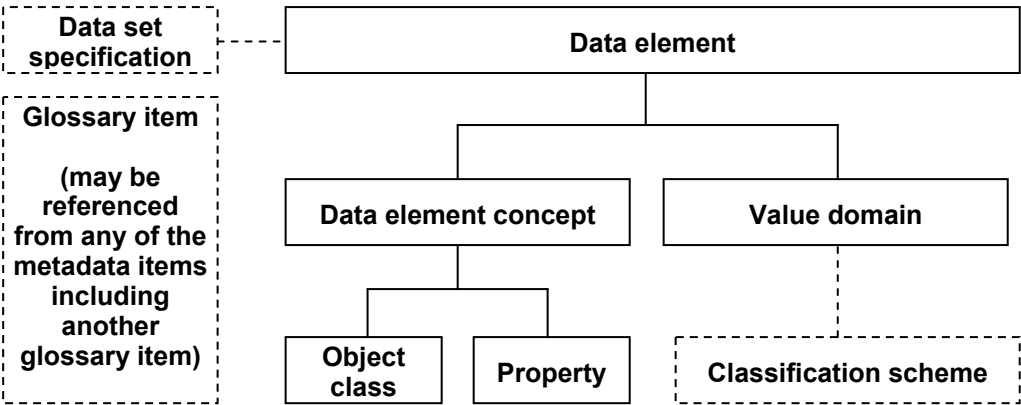


Figure 1: Types of data standards described in this dictionary and METeOR

Data element

A data element is the basic unit of identifiable and definable information created by combining a data element concept and a value domain. In the dictionary, examples of data elements include *Person – date of birth, DDMMYYYY*, and *Referral – contact method, code N*.

Data elements are used to standardise the meaning and representation of data in order to ensure consistency of the data collected.

The conceptual meaning of the data element is captured within the data element concept. The representation of the data element is captured within the value domain. A data element

is composed of only one data element concept and only one value domain, that is, a data element is the union of one data element concept and one value domain.

The union of a specific data element concept with a specific value domain creates a data element. For example the *Person – date of birth* example mentioned above can be combined with the *Date DDMMYYYY* value domain to create the data element: *Person – date of birth, DDMMYYYY*. Alternatively, the data element could be formed using the *Date YYYY* value domain making a distinct data element *Person – date of birth, YYYY*.

Data element concept

A data element concept is defined as a concept created for the purposes of defining a data element by the union of an object class and a property. Only one object class and one property can be joined for each data element concept. Within the dictionary, examples of data element concepts include *Person – date of birth* and *Service provider organisation – service delivery setting*.

The data element concept represents the concept for which data are sought but not how the data are to be collected. The specification of how the data should be collected is achieved through the union with a value domain.

Value domain

A value domain specifies the representation of the data element. It specifies how something is to be represented. A value domain can specify any of the following:

- the range of permitted values, for example, a measure of weight in grams represented by 3 numeric characters
- all permissible values as a set of codes, for example, Code 1 Female, Code 2 Male
- referencing the values documented in a nationally or internationally endorsed classification, such as all codes in the Australian Standard Classification of Languages 2005, or all activity codes listed in the National Classifications of Community Services (NCCS V2 2002)⁴.

Classification scheme

In the dictionary, a classification scheme is defined as an official terminological system, recognised and endorsed by a national or international body, that is used to classify data.

Examples of classification schemes include:

- Australian Standard Classification of Education 2001
- Australian Standard Classification of Languages 2005
- Australian Standard Classification of Occupations, 2nd edition.

⁴ Australian Institute of Health and Welfare 2003. *National classifications of community services, version 2.0*. AIHW cat. no. HWI 40. Canberra: Australian Institute of Health and Welfare.

A classification scheme may be implemented in one or more data elements. For example, the Australian Standard Classification of Languages 2005 is implemented in the data elements *Person – preferred language, code (ASCL 2005) NN{NN}*, *Person – main language other than English spoken at home, code (ASCL 2005) NN{NN}*, and *Person – first language spoken, code (ASCL 2005) NN{NN}*.

Object class

An object class describes the entity (the ‘thing’) about which we want to collect data. A property describes the particular characteristic or attribute of that entity. As shown in Figure 1, the union of one object class with one property specifies what is to be described.

Within the dictionary examples of object classes include *Person*, *Service provider organisation*, and *Service episode*.

Object classes can be specialisations of other object classes. Specialisations allow object classes to be grouped and subtyped in a meaningful manner and help users browse through and locate relevant object classes. In a specialisation tree, an object class can only be associated with a single parent object class but may have more than one child object class. A child object class inherits all characteristics of its parent object class, but a child object class may have unique characteristics. Depending on what information we want to know, broadly defined object classes (such as *Person*, *Service provider organisation* or *Service event*) or more specific object classes (such as *Client* or *Child*) may be used.

Property

A property is a characteristic of an object class of interest. For example, the object class *Person* can have characteristics such as sex and date of birth. These characteristics are referred to as properties.

The combination of a specific property with a specific object class creates a data element concept. The above mentioned examples create the data element concepts *Person – sex* and *Person – date of birth*.

Properties are assigned property groups which group similar properties such as lifestyle characteristics and financial characteristics. These property groups help users browse through and locate relevant properties.

Glossary item

A glossary item defines the meaning of a term within a specific context. Within METeOR examples of glossary items include ‘Adoption’ and ‘Family’. These things of interest are not currently defined as object classes but their meaning must be understood for data to be collected. A glossary item can be linked to any other metadata item type. For example, the data element definition for *Geographic location of person* contains a link to the glossary item ‘Statistical Local Area (SLA)’. Glossary items are not specified in ISO/IEC 11179.

Data set specification

Data elements may be grouped for a specific purpose within a data set specification. A data set specification is either prescribed for national collection and reporting as a national minimum data set (NMDS) or endorsed as the recommended collection in a data set specification (DSS).

Metadata attributes

The template for describing a data element is described in Appendix 2.

Appendix 2: Data element template

Common name, for example, 'Height-measured'

Identifying and definitional attributes

Metadata item type	The name of the type of data standard, that is: Data element
Technical name	The name of the type of data standard following strict naming conventions, for example: Person – height (measured), total centimetres NN[N].N
Synonymous name	A synonym or list of synonyms for the name within the specified context. This attribute may be left blank.
METeOR identifier	A unique identifier within METeOR, for example: 270361
Registration status	A status value for the data standard indicating its stage in the registration process, for example: NCSIMG, Standard 01/03/2005 Explanation: NCSIMG approved this data element as a national standard on 1 March 2005.
Definition	A concise statement that expresses the essential nature of the data standard and its differentiation from other data standards, for example: a person's measured height
Context	A designation and/or description of the application, environment or discipline in which the definition of the data element is valid. This attribute may be left blank.
Data element concept	The data element concept implemented in this data element, for example: Person – height

Explanation: This data element describes the union of the object class 'Person' and the property 'Height'.

Value domain

The value domain implemented in this data element, for example:

Total centimetres NN[N].N

Explanation: This data element is a measurement in centimetres, which accepts numeric values in the format of NN.N or NNN.N.

Collection and usage attributes

Guide for use

Comments, advice or instructions for the interpretation or application of the data standard, for example:

To ensure consistency in measurement, the measurement protocol described under 'Collection methods' should be used.

Collection methods

Comments, advice or instructions for the actual capture of data, for example:

All equipment should be checked before each measurement to ensure that both the headboard and floor are at 90 degrees to the vertical rule.

Comments

Any additional information that adds to the understanding of the data standard, for example:

It may be desirable to present height data in categories. It is recommended that 5 cm groupings are used for this purpose.

Source and reference attributes

Submitting organisation

The full name of each organisation responsible for the submission of the data standard for endorsement as a national standard, for example:

Australian Institute of Health and Welfare

Steward

The full name of the organisation that has accepted responsibility and been approved by a registration authority to provide ongoing maintenance and management of a data standard, for example:

Australian Institute of Health and Welfare

Origin

Any document(s) (including websites), organisation(s) or committee(s) from which any content of the metadata item

originates, using AIHW referencing guidelines.

This attribute may be left blank.

Reference documents

The reference to any document that contributed to the development of the data standard which were not cited above, using AIHW referencing guidelines*.

Relational attributes

Related metadata

An indicator of relationships between data standards within a given sector (health, community services or housing assistance), for example:

Supersedes Height – measured, version 2, DE, NHDD, NCSIMG, Superseded 01/03/2005.pdf (28.7 KB)
NCSIMG, Standard 01/03/2005

Is used in the formation of 'Adult – body mass index (measured)', ratio NN[N].N[N] NCSIMG, Standard 01/03/2005

Explanation: This data element replaced the old version of the data element as of 1 March 2005. It is used in the calculation of body mass index, which was also approved as a standard on the same date.

Appendix 3: Guide to data development

Data development is the process of building a data set for a specific purpose. This appendix outlines some data development principles. Further information on data development may be obtained from the AIHW's National Data Development and Standards Unit.

Data development principles

Creating data standards is part of data development

The quality of data, including its consistency and comparability, is enhanced when data standards are available to support the collection and use of a data set. The development of data standards is not done at the end of the data development process. It is very much part of the data development process and carries on throughout the life of the data set.

Data developers need to develop operational procedures for making data standards a key component of their data development process and for ensuring their ongoing relevance and maintenance.

There is a cost associated with creating data standards. However, the cost of not creating data standards may be even higher. This includes loss of information with staff changes; data redundancy; data conflicts; liability; misapplications; and decisions based on poorly documented data. These costs should be factored into the data development budget.

National and international standards should be used wherever available and applicable

When data are being developed, it is important to ensure that the specifications of the data are consistent with national and international data standards. This helps to avoid the duplication of effort, and development of conflicting data standards.

Data sets should be based on a single set of agreed definitions and data standards, for example the *National community services data dictionary*. This allows data developers to pick and mix from existing agreed definitions, ensuring a high degree of consistency and reducing data development time and cost.

Data development is system independent

Data development must not be limited by the capability of any particular system. Data development must ensure that data in the data set are well defined and standardised to be comparable independent of the organisation, system or tool that captures the data.

Data must be fit for the intended purpose

Data must be fit for the intended purpose and use. Data development should define the purpose of the data to be collected, and where limitations of the data exist, these need to be acknowledged.

Although it is important to be mindful of the opportunities to reuse data where appropriate, it is also important to note that data suitable for some purposes may be limited in use for others. For example, data about 'presenting problem' and 'reason for encounter' may be useful for evaluating hospital emergency department services, treating patients and managing department resources, but may not be adequate for informing about the incidence of domestic violence or child abuse.

Data development may be incremental

Data development should support incremental development of data, such that the scope of the data set is expanded over time.

It may not be possible to develop all data required for a data set at the same time. Some data may be more readily agreed on and easily collected. Other data may be more problematic and require more time to develop. With a data set about problem gambling, for example, it may be quite easy to develop demographic data about clients seeking help in relation to their gambling problems, but more difficult to develop data that can be used to measure outcomes of relevant education programs or assistance services.

Data development should be mindful of privacy concerns

Data development processes must take account of information privacy principles and security policies. Data sets should avoid the inclusion of data that may be regarded as private or confidential in nature; otherwise, data may not be reliable or accurate because of respondents' reluctance to provide such information.

Data development should minimise collector/recording burden

Good data development should ensure compatibility of data collection and reporting requirements to avoid situations where the same data have to be collected (or counted) or reported differently for different programs. This will reduce the reporting burden on service providers and help to reduce the cost of data development.

Data development should reflect not drive practice

Data developed must be data that can be reasonably expected to be collected as a by-product of service delivery or administrative practice. The data must be relevant and meaningful to those collecting the data and be of benefit to service providers. Data development must take account of business needs, feasibility of data collection and appropriateness of the data, as well as scientific evidence and recommendations (guidelines) of subject matter experts. Where possible, data development should be based on data that service providers already want or need to collect about clients and service provision. Most importantly, data development should ensure that data collectors are not forced to operate differently from their usual practice. For example, data about a client's country of birth is usually not relevant

for service delivery and should not be part of routine data set collections. Of more relevance to service delivery is data about whether interpreter services are required and the client's preferred language. Data about country of birth, which may be useful for statistical purposes, can be collected as part of 'one-off' surveys of clients.

Create once, use often

An important principle of data development is that data needed to support secondary (or downstream) information purposes (for example reporting, policy, governance, decision support) should be derivable from primary data (point of service delivery data). Otherwise, data required for downstream requirements will be developed and collected separately, resulting in significant additional costs because of the need to establish parallel data collection systems to support existing and new data flows, whose products may not necessarily be integratable. Similarly data developed and collected for mainly statistical purposes should be used to provide feedback to improve and enhance primary service delivery.

Appendix 4: The National Community Services Data Committee

The National Community Services Data Committee is a subcommittee of the National Community Services Information Management Group. It was mainly established to develop and maintain the *National community services data dictionary* and minimum data sets in all areas of community services. The National Community Services Data Committee has a coordinating role to ensure national consistency of data definitions and standards and in quality control.

The functions of the National Community Services Data Committee are to:

- be responsible for overseeing the development and maintenance of the *National community services data dictionary* and promoting consistency between its standards and definitions and those in the *National health data dictionary* and the *National housing assistance data dictionary*
- receive, consider and comment on data definitions, NMDS, and the collection of data items, and make recommendations to the Management Group for endorsement of their inclusion in the *National community services data dictionary*, and the *Metadata online registry*
- with advice from the Management Group on the national priorities and work program, produce a work plan for approval by the Management Group and report on progress of each working party twice a year to the Management Group
- actively seek out data definition activities to inform the National Community Services Information Work Program of information developments that meet (or have the potential to meet) specified criteria for inclusion on the work program
- develop links and foster cooperative working arrangements within the community services sector and between other sectors on data development activities
- document relevant current and planned data development activities in each jurisdiction
- develop and maintain processes and guidelines for the development of national data standards
- develop and maintain national processes and guidelines for disseminating data definitions and standards to data collection agencies
- take the lead role in the development of national community services definitions.

Table A1: National Community Services Data Committee membership at May 2006

Name/Jurisdiction	Position/organisation	Contact details
Anne Jenkins Chair	Manager, Data and Research Department of Disability, Housing and Community Services ACT Government GPO Box 158, CANBERRA ACT 2601	Ph: 02 6205-0082 Fax: 02 6205-0343 Email: anne.jenkins@act.gov.au
Tony Carr Deputy Chair	Manager Quality and Statistics Office for Children 555 Collins Street MELBOURNE VIC 3000	Ph: 03 9616 7225 Fax: 03 9616 8887 Email: tony.carr@dhs.vic.gov.au
David Martin Department of Health and Ageing	Manager Management Information and Data Analysis Section GPO Box 9848 CANBERRA ACT 2601	Ph: 02 6289 5095 Fax: 02 6289 1267 Email: david.martin@health.gov.au
Sean Terry Tasmania	Manager Planning, Performance and Evaluation Strategic Services Department of Health and Human Services GPO Box 125 HOBART TAS 7001	Ph: 03 6233 6720 Fax: 03 62334843 Email: sean.terry@dhhs.tas.gov.au
Adam Tonkin Centrelink	Project Manager Business Definitions Registry Data Services Branch PO Box 7788 CANBERRA MAIL CENTRE ACT 2610	Ph: 02 6244 6577 Email: adam.tonkin@centrelink.gov.au
John Sant Australian Bureau of Statistics	Population Statistics Standards Australian Bureau of Statistics Locked Bag 10 BELCONNEN ACT 2616	Ph: 02 6252 5717 Fax: 02 6252 5281 Email: john.sant@abs.gov.au
Meera Rajendran Secretariat	Acting Head National Data Development and Standards Unit GPO Box 570 CANBERRA ACT 2601	Ph: 6244 1195 Fax: 02 6244 1069 Email: meera.rajendran@aihw.gov.au
Diane Gibson Australian Institute of Health and Welfare	Head, Welfare Division Australian Institute of Health and Welfare GPO Box 570 CANBERRA ACT 2601	Ph: 02 6244 1190 Email: diane.gibson@aihw.gov.au
Susan Tregeagle Barnardos	Senior Manager Program Services GPO Box 9996 SYDNEY NSW 2001	Ph: 02 6161 6264 Ph: 02 9881 7933 Fax: 02 6161 6365 Fax: 02 9281 0441 Email: suetreg@barnardos.org.au

(continued)

Table A1 (continued): National Community Services Data Committee membership at May 2006

Name/Jurisdiction	Position/organisation	Contact details
Allison Shoobridge Commonwealth Grants Commission	Expenditure Branch Cypress Court 5 Torrens Street CANBERRA ACT 2612	Ph: 02 6229 8817 Fax: 02 6229 8817 Email: alison.shoobrige@cgc.gov.au
Greg Nockolds Department for Community Development, WA	Data Architect, Information Planning Policy Standards and Applications Branch Information Services Division PO Box 6334 EAST PERTH WA 6892	Ph: 08 9222 2770 Fax: 08 9222 2942 Email: gregn@dcd.wa.gov.au
Arthur Hume Department for Families and Communities, SA	Senior Project Officer Research and Analysis Unit GPO Box 292 ADELAIDE SA 5001	Ph: 08 8226 6747 Fax: 08 8226 6626 Email: arthur.hume@dfc.sa.gov.au
Proshanta Dey Department of Community Services, NSW	Senior Project Officer Information Management 4-6 Cavill Ave ASHFIELD NSW 2131	Ph: 02 9716 2212 Fax: 02 9716 2290 Email: proshanta.dey@community.nsw.gov.au
Richard Lennon Department of Communities, Qld	A/Assistant Director Strategic Planning and Performance Measurement Branch Corporate and Executive Services Directorate GPO Box 806 Brisbane QLD 4000	Ph: 07 3224 5831 Fax: 07 3224 7914 richard.lennon@communities.qld.gov.au
Phong Lamuri Territory Health	87 Mitchell Street DARWIN NT 0801	Ph: 08 8999 2835 Email: phong.lamuri@nt.gov.au
Teri Jenkins Department of Families, Community Services and Indigenous Affairs	Teri Jenkins Section Manager Data Strategy Information Management Branch Box 7788 CANBERRA ACT 2610	Ph: (02) 6244 6135 Fax: 02 6244 7981 Email: teri.jenkins@facs.gov.au
Peter Munday Lifeline Australia Inc.	Justlook Coordinatorr GPO Box 583 CANBERRA ACT 2601	Ph: 02 6215 9413 Fax: 02 6282 6566 Email: peterm@lifeline.org.au
Eleri Morgan-Thomas Mission Australia	National Manager Community Services PO Box 3515 SYDNEY NSW 2001	Ph: 02 9212 1116 Email: morgan-thomas@missionaustralia.com.au
Keith Suter Wesley Mission	Consultant for Strategic Planning PO Box A5555 SYDNEY SOUTH NSW 1235	Ph: 02 9263 5400 Fax: 02 9282 1140 Email: kdsuter@wesleymission.org.au
Tanya Wordsworth Secretariat	National Data Development and Standards Unit Australian Institute of Health and Welfare GPO Box 570 CANBERRA ACT 2601	Ph: 02 6244 1254 Fax: 02 6244 1069 Email: tanya.wordsworth@aihw.gov.au