A guide to data development

National Data Development and Standards Unit

2007

Australian Institute of Health and Welfare
Canberra
AIHW cat. no. HWI 94
Contents

Acknowledgments................................................................................................................ .............. vi
Abbreviations.................................................................................................................. ...................vii
Summary ........................................................................................................................................... viii

1 Introduction................................................................................................................. .......................1
   1.1 Purpose.................................................................................................................... ..................1
   1.2 Audience ...................................................................................................................2
   1.3 Stakeholders ..............................................................................................................2
   1.4 Structure of the guide............................................................................................. .........3

2 Data development overview .................................................................................................... .......4
   2.1 What is data? ............................................................................................................................4
   2.2 What is information?....................................................................................................... ........4
       2.2.1 Uses of information...................................................................................................... .4
   2.3 What is metadata? .......................................................................................................... .........5
       2.3.1 Metadata is the key to data..........................................................................................5
       2.3.2 Uses of metadata ...........................................................................................................6
   2.4 What is ISO/IEC 11179? ..................................................................................................... ....7
       2.4.1 Parts of ISO/IEC 11179 ................................................................................................7
   2.5 What is a data element? ..........................................................................................................8
   2.6 What is a metadata registry?............................................................................................... ...9
   2.7 What is data development?.................................................................................................. 10
       2.7.1 Data development for primary and secondary purposes .....................................11
   2.8 What is a data set? .................................................................................................................12
       2.8.1 Data set specifications ................................................................................................12
   2.9 Data development and metadata ........................................................................................13
   2.10 Data development and data standards.............................................................................13
       2.10.1 Why standardise data?.............................................................................................14
   2.11 What are national data standards?....................................................................................14
   2.12 Terminology .........................................................................................................................15
   2.13 Data development and data quality..................................................................................15

3 Data development principles and methodology.......................................................................17
3.1 Principles of good data development ................................................................. 17
  3.1.1 Creating data standards is part of data development ................................. 17
  3.1.2 National and international standards should be used wherever available and applicable ................................................................. 17
  3.1.3 Be clear about the purpose of the data collection ........................................... 18
  3.1.4 Data included must be required to meet the objectives of the data collection... 18
  3.1.5 Create once, use often ............................................................................... 18
  3.1.6 Acknowledge the limitations of data ............................................................ 18
  3.1.7 Data development may be incremental ......................................................... 19
  3.1.8 Data development is system independent ...................................................... 19
  3.1.9 Data development should be mindful of privacy concerns ............................ 19
  3.1.10 Data development should minimise collector/recording burden ................. 19
  3.1.11 Data development should reflect not drive practice .................................... 19
3.2 The data development process ........................................................................ 20
  3.2.1 Business context and information need identification ................................. 20
  3.2.2 Feasibility analysis ....................................................................................... 20
  3.2.3 Consultation and collaboration ................................................................... 23
  3.2.4 Identifying data for development ................................................................. 25
  3.2.5 Developing data elements ........................................................................... 27
  3.2.6 Field testing .................................................................................................. 29
  3.2.7 Estimating cost of implementation ............................................................... 29
  3.2.8 Obtaining authoritative endorsement .......................................................... 30
3.3 Risk management .............................................................................................. 31
4 Developing data elements .................................................................................... 32
  4.1 The components of a data element ................................................................. 32
    4.1.1 Object Class ............................................................................................... 33
    4.1.2 Property ..................................................................................................... 34
    4.1.3 Data Element Concept (DEC) ................................................................. 35
    4.1.4 Value Domain ......................................................................................... 36
    4.1.5 Classification ......................................................................................... 38
  4.2 Formulating data elements .............................................................................. 39
    4.2.1 Formulating data elements – an example ................................................. 40
  4.3 Guidelines for good data standards ................................................................. 41
    4.3.1 Name – guidelines .................................................................................. 41
    4.3.2 Definition – guidelines .......................................................................... 42
4.3.3 Context – guidelines .................................................................46
4.3.4 Value domain – guidelines ........................................................47
4.3.5 Permissible values – guidelines ..................................................49
4.3.6 Guide for use – guidelines .........................................................53
4.3.7 Collection methods – guidelines ................................................53

5 METeOR .................................................................................................55
  5.1 What is METeOR? ...........................................................................55
  5.2 Principles of METeOR .................................................................56
  5.3 METeOR metadata item types and attributes ...............................57

6 Agreements and governance arrangements for national data standards ................................72
  6.1 National Health Information Agreement (NHIA) ..........................72
  6.2 National health information governance arrangements ...............73
  6.3 National Community Services Information Agreement (NCSIA) ....75
  6.4 Community services information governance arrangements .......75
  6.5 National Housing Data Agreement (NHDA) .................................77
  6.6 Agreement on National Indigenous Housing Information (ANIHI) .77
  6.7 Housing information governance arrangements ..........................77
  6.8 States and territory governance arrangements ............................79
  6.9 Local governance arrangements ..................................................79

7 Who to contact for help ........................................................................81

List of tables ............................................................................................83
List of figures ...........................................................................................84
References ...............................................................................................85
Acknowledgments

The Guide was written by Meera Rajendran of the National Data Development and Standards Unit (NDDSU) of the Australian Institute of Health and Welfare, who would like to acknowledge the contribution of the staff of the NDDSU, especially David Neilsen who documented the guidelines for formulating good data standards, and Shari Beaumont for formatting and proofreading the document.

The author would also like to thank the New South Wales Health Department, the Queensland Health Department and Standards Australia IT–14–2, all of whom contributed to the Guide.

Thanks are due to Trish Ryan, former Unit Head, NDDSU, whose idea it was in the first place to write this Guide, so that the Institute’s wealth of knowledge and experience about data development could be made available, and contribute towards the development of good data collections.
# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABS</td>
<td>Australian Bureau of Statistics</td>
</tr>
<tr>
<td>AHMAC</td>
<td>Australian Health Ministers Advisory Council</td>
</tr>
<tr>
<td>AIHW</td>
<td>Australian Institute of Health and Welfare</td>
</tr>
<tr>
<td>ANIHI</td>
<td>Agreement on National Indigenous Housing Information</td>
</tr>
<tr>
<td>CSMAC</td>
<td>Community Services Ministers’ Advisory Council</td>
</tr>
<tr>
<td>DEC</td>
<td>Data element concept</td>
</tr>
<tr>
<td>DSS</td>
<td>Data set specification</td>
</tr>
<tr>
<td>HACC</td>
<td>Housing and Community Care</td>
</tr>
<tr>
<td>HDSC</td>
<td>Health Data Standards Committee</td>
</tr>
<tr>
<td>HMAC</td>
<td>Housing Minister’s Advisory Council</td>
</tr>
<tr>
<td>MDS</td>
<td>Minimum data set</td>
</tr>
<tr>
<td>METeOR</td>
<td>Metadata Online Registry</td>
</tr>
<tr>
<td>NCHI</td>
<td>National Committee for Housing Information</td>
</tr>
<tr>
<td>NCSDC</td>
<td>National Community Services Data Committee</td>
</tr>
<tr>
<td>NCSDDD</td>
<td>National Community Services Data Dictionary</td>
</tr>
<tr>
<td>NCSIA</td>
<td>National Community Services Information Agreement</td>
</tr>
<tr>
<td>NCSIMG</td>
<td>National Community Services Information Management Group</td>
</tr>
<tr>
<td>NDDSU</td>
<td>National Data Development and Standards Unit</td>
</tr>
<tr>
<td>NEHTA</td>
<td>National e-Health Transition Authority</td>
</tr>
<tr>
<td>NHADD</td>
<td>National Housing Assistance Data Dictionary</td>
</tr>
<tr>
<td>NHDA</td>
<td>National Housing Data Agreement</td>
</tr>
<tr>
<td>NHDD</td>
<td>National Health Data Dictionary</td>
</tr>
<tr>
<td>NHIA</td>
<td>National Health Information Agreement</td>
</tr>
<tr>
<td>NHIMPC</td>
<td>National Health Information Management Principal Committee</td>
</tr>
<tr>
<td>NMDS</td>
<td>National minimum data set</td>
</tr>
<tr>
<td>SIMC</td>
<td>Statistical Information Management Committee</td>
</tr>
</tbody>
</table>
Summary

For more than a decade, the Australian Institute of Health and Welfare (AIHW) has been undertaking data development and assisting other organisations with related activities. In doing so, the AIHW has developed a body of expertise in data development. Much of this experience has been incorporated into this guide.

While data development activities are undertaken by many organisations, detailed information about the process of data development is not readily available. Data development requires a considerable investment of time, staff and other resources, and includes the development of data standards to support consistent data collection. The guide outlines the basics of data development and the steps to follow when developing data.

The use of electronic health records is increasing throughout the health care system and the feasibility of electronic continuous client records is being considered by the community care sector. The advent of electronic data capture and communication systems provides a unique opportunity to improve the quality and coverage of information in Australia and the potential to re-use and share data becomes even more relevant. In line with the ‘create once, use often’ objective of information management, it is important to look for opportunities to improve the information available for national reporting and analysis when undertaking electronic data collection initiatives. Otherwise, significant additional costs may be incurred as a result of the need to create parallel systems whose products may not necessarily be compatible, resulting in inefficiencies and risks to data quality. It is therefore important that data development be undertaken using principles that ensure data collected at the point of service delivery can be aggregated and used for secondary purposes.

The guide outlines sound data development practices that incorporate the necessary steps to produce high-quality, consistent data that meets the needs of all users of the data, and attracts the necessary level of agreement on content and authority behind the data specifications. The document promotes the role of metadata in describing data and the use of data standards to improve consistency in data recording where data consistency and comparability is required. The document takes the reader from the basics of data development to the methodology behind good data development and the specification of data elements to describe and standardise data.

The principles and processes outlined are applicable to all subject areas where data are required, including community services, health (including clinical data collections) and housing assistance. The processes outlined are relevant for local data collections as well as state/territory or national data collections. The document provides information about governance processes for national data collections.