Guidelines for documentation of species and vegetation data

(from Environment Australia: http://www.deh.gov.au/erin/documentation/metadata.html)

These guidelines are modified from the ANZLIC Metadata Guidelines *Version 1, July 1996. The full ANZLIC document is available through the Internet at:* **<http://www.anzlic.org.au/metaelem.htm>**

Category	Element	Comment	Page
Data	Title	The ordinary name of the data.	**2**
	Custodian	The organisation responsible for the data.	**2**
Description	Abstract	A short description of the contents of the data. This must include details of scale or cell size. Scale is the ratio or fraction between the distance on a map chart or photograph and the corresponding distance in the real world. Cell size is the dimension of grid cells or pixels. Environment Australia requires a list of the attributes included in the data, along with a brief description of each. This list is critical for assessing what information is actually contained within the data.	**3**
	Search Word(s)	Words likely to be used by a non-expert to look for the data.	**4**
	Geographic Extent Name(s) **OR**	A list of pre defined geographic extents (such as map sheets, local government areas, catchments, States) that reasonably indicate the spatial coverage of the data.	**6**
	Geographic Extent Polygon(s)	An alternate way of describing geographic extent if no pre-defined area is satisfactory.	**7**
Data Currency	Beginning date	Earliest date of records in the data.	**8**
	Ending date	Last date of information in the data.	**8**
Data Status	Progress	The status of the process of creation of the data.	**8**
	Maintenance and Update Frequency	Frequency of changes or additions made to the data.	**9**
Access	Access Constraint	Any restrictions or legal prerequisites applying to the use of the data, eg. licence.	**9**
Data Quality	Lineage	A brief history of the source and processing steps used to produce the data. Lineage may also include details of scale or accuracy (see also under the Abstract Element).	**10**
	Positional Accuracy	A brief assessment of the closeness of the location of spatial objects in the data in relation to their true position on the Earth.	**11**
	Attribute Accuracy	A brief assessment of the reliability assigned to features in the data in relation to their real world values. This element should include a list of attributes and information on the accuracy of each.	**12**
	Logical Consistency	A brief assessment of the logical relationships between items in the data.	**13**
	Completeness	A brief assessment of the completeness of coverage, classification and verification.	**14**
Contact Information	Contact Organisation	Ordinary name of the organisation from which the data may be obtained.	**15**
	Contact Position	The relevant position in the Contact Organisation.	**15**
	Mail Address	Postal address of the Contact Position.	**15**
	Suburb or Place or Locality	Suburb of the Mail Address.	**15**
	State	State of Mail Address.	**15**
	Postcode	Postcode of the Mail Address.	**15**
	Telephone	Telephone of the Contact Position.	**15**
	Facsimile	Facsimile of the Contact Position.	**15**
	Electronic Mail Address	Electronic Mail Address of the Contact Position.	**15**
Additional Metadata	Additional Metadata	Reference to other directories or systems containing further information about the data.	**15**

Element: TITLE

This is the name given to the data by the custodial organisation. It should convey a clear impression of the information contained in the data and should be in plain language (i.e. preferably not solely in acronym form).

A dataset is any collection of information that can be geographically referenced. It may be about environment, transport, population, utilities, property or any other geographic information. The data need not be in digital form.

The objective of this element is to provide a unique reference to the data.

Allowable Content:

The title of the dataset should be easily understood by a general user. In other words, it should not be composed entirely of acronyms or short titles—the meaning of which would only be evident to an existing user or someone in the custodian organisation associated with the data. The title should be as descriptive as possible.

If the data are usually referred to via an acronym, this should appear in brackets at the end of the title, following a descriptive rendering of the title.

Example(s):

Example 1:	Endangered and vulnerable vertebrate specimen records from the Australian Museum
Example 2:	A study of the flora of the Clarence Valley
Example 3:	A vegetation survey of the Murwillumbah 1:250,000 map sheet.

Element: CUSTODIAN

Definition of Element:

Use the ordinary business name of the custodian organisation for the dataset. Custodianship is a concept adopted by ANZLIC and which has been described in more detail in ANZLIC Issues Paper No 1, April 1990.

The custodian is the organisation responsible for ensuring the accuracy, currency, storage, security and distribution of the dataset. In fulfilling these responsibilities, the custodian is expected to consult with, and take into account the needs of users other than itself. The custodian may choose to delegate these functions while still retaining responsibility.

The custodian of a dataset need not necessarily be the holder of the copyright, or the originator of the data, although in many cases the custodian will be both of these.

A custodian is not necessarily a government agency. A private organisation or an individual may be a custodian. However, in many cases the custodian will be a government organisation, reflecting the fact that in Australia and New Zealand a vast amount of geographic data is created and managed in government.

Allowable Content:

Use the ordinary business name of the organisation in full.

If the organisation is also known by a common acronym, please include this at the end of the full name. If the organisation is usually known only by an acronym, use only the acronym without brackets.

Example(s)

Example 1:	Australian Museum
Example 2	Queensland Herbarium
Example 3:	CSIRO Wildlife and Ecology
Example 4:	Australian Surveying and Land Information Group (AUSLIG)
Example 4:	Australian Surveying and Land Information Group (AUSLIG)

Element: ABSTRACT

Definition of Element:

A characterisation of the data and its contents. It may be a brief narrative, a summary or an abstract and should be in plain English with Acronyms spelled out. Details of scale **or** cell size ***must*** be included.

Scale or Cell Size:

Scale:
Ratio or fraction between the distance on a map, chart or photograph and the corresponding distance in the real world. Scale is usually used if the data were derived by digitising from printed maps or aerial photographs. Scale reflects the resolution of the data. Larger scale data is preferable where available and of high quality.

OR

Cell Size:
Dimension of grid cells or pixels. This field is used when describing raster or gridded data. The cell size affects the spatial accuracy of the data and any derived products. The smaller the grid cell, the higher the resolution of the data.

NB. *For point-based specimen data from an herbarium or museum where the data may be from a range of sources and be highly variable in its accuracy, Scale or Cell size may not be applicable. This should be noted in the metadata under* ***Spatial Accuracy***.

Other material to be conveyed in this element may include a description of the purpose for which the data were created, and a textual description of the spatial extent of the data contained in the data.

For vegetation data, attributes such as whether pre-1750 or extant/present vegetation, the type of mapping (ie. floristic, structural or both) and details of the survey methods including plot size, frequency of survey, disturbance attributes and sampling strategy, should be described.

Allowable Content:

The Abstract should describe the contents of the data for a non expert user, in plain language. The field is a maximum of 2000 characters.

Example(s):

Example 1: [Plant specimen records from the National Herbarium of NSW]

Specimen data for major landcover species, threatened species and various other taxa from the NSW National Herbarium. Attributes include Name (genus, species, infraspecies rank, infraspecies name), Latitude, Longitude, Locality, Altitude, Collector name and number, Habitat, Institution name and number and Date. Point records with no scale.

Example 2: [Beecroft Peninsula Vegetation Survey]

The report describes the present vegetation of the Royal Australia Navy Weapons Range - Beecroft at 1:250,000 scale. The report provides managers of the Range with a tool for identifying floral communities, to assess conservation values of locations on the Range and effectively allocate management resources. 8 primary and 27 secondary different vegetation communities are identified and described in detail. Twinspan was used as the analytical tool to classify about 140 sites and 573 species. Only 25 species were weeds. Estimates of fire frequency were determined from the number of branches on *Banksia* species.

Example 3: [Vegetation Survey data: Coongie Biological Survey, South Australia]

The survey area is in the Far northeast of South Australia and covers the Coongie Lakes Region. The survey documents the distribution and abundance of vegetation and vertebrate species in the study area at the time of the survey. The survey used a standard 100-meter grid size.

Element: SEARCH WORD(S)

Definition of Element:

Words likely to be used for searching by a person who does not necessarily have expertise in the subject matter being searched. The words are created from the user or searchers viewpoint not from that of the writer of the abstract or the Custodian of the data. Where a number of terms may be relevant to the content of the data, the most concise term should be used. When a dataset contains diverse information, multiple search words may be allocated.

Allowable Content:

AGRICULTURE
AGRICULTURE Crops
AGRICULTURE Livestock
AGRICULTURE Horticulture
AGRICULTURE Irrigation

ATMOSPHERE
ATMOSPHERE Air Quality
ATMOSPHERE Ozone
ATMOSPHERE Greenhouse
ATMOSPHERE Pressure

BOUNDARIES
BOUNDARIES Administrative
BOUNDARIES Biophysical
BOUNDARIES Cultural

CLIMATE AND WEATHER
CLIMATE AND WEATHER Meteorology
CLIMATE AND WEATHER Climate change
CLIMATE AND WEATHER Drought
CLIMATE AND WEATHER El Nino
CLIMATE AND WEATHER Extreme weather events
CLIMATE AND WEATHER Radiation
CLIMATE AND WEATHER Rainfall
CLIMATE AND WEATHER Temperature

DEMOGRAPHY

DISEASE

ECOLOGY
ECOLOGY Community
ECOLOGY Ecosystem
ECOLOGY Habitat
ECOLOGY Landscape

ENERGY
ENERGY Coal
ENERGY Electricity
ENERGY Petroleum
ENERGY Renewable
ENERGY Use

FAUNA
FAUNA Exotic
FAUNA Insects

FAUNA Invertebrates
FAUNA Native
FAUNA Vertebrates

INDUSTRY Mining
INDUSTRY Primary
INDUSTRY Service

FISHERIES
FISHERIES Aquaculture
FISHERIES Freshwater
FISHERIES Marine
FISHERIES Recreational

FLORA
FLORA Exotic
FLORA Native

FORESTS
FORESTS Agriforestry
FORESTS Natural
FORESTS Plantation

GEOSCIENCES
GEOSCIENCES Hydrogeology
GEOSCIENCES Geochemistry
GEOSCIENCES Geology
GEOSCIENCES Geomorphology
GEOSCIENCES Geophysics

HAZARDS
HAZARDS Cyclones
HAZARDS Drought
HAZARDS Earthquake
HAZARDS Fire
HAZARDS Flood
HAZARDS Landslip
HAZARDS Manmade
HAZARDS Pests
HAZARDS Severe local storms
HAZARDS Tsunamis

HEALTH

HERITAGE
HERITAGE Aboriginal
HERITAGE Architectural
HERITAGE Natural
HERITAGE World

HUMAN ENVIRONMENT
HUMAN ENVIRONMENT Economics
HUMAN ENVIRONMENT Housing
HUMAN ENVIRONMENT Livability
HUMAN ENVIRONMENT Planning
HUMAN ENVIRONMENT Structures and Facilities
HUMAN ENVIRONMENT Urban Design

INDUSTRY
INDUSTRY Manufacturing
POLLUTION Soil
POLLUTION Water

INDUSTRY Other

LAND
LAND Cadastre
LAND Cover
LAND Geodesy
LAND Geography
LAND Ownership
LAND Topography
LAND Use
LAND Valuation

MARINE
MARINE Biology
MARINE Coasts
MARINE Estuaries
MARINE Geology and Geophysics
MARINE Reefs
MARINE Human Impacts
MARINE Meteorology

MINERALS

MOLECULAR BIOLOGY
MOLECULAR BIOLOGY Genetics

OCEANOGRAPHY
OCEANOGRAPHY Physical
OCEANOGRAPHY Chemical

PHOTOGRAPHY AND IMAGERY
PHOTOGRAPHY AND IMAGERY Aerial
PHOTOGRAPHY AND IMAGERY Remote Sensing
PHOTOGRAPHY AND IMAGERY Satellite

POLLUTION
POLLUTION Air
POLLUTION Noise

SOIL
SOIL Erosion
SOIL Biology
SOIL Chemistry
SOIL Physics

TRANSPORTATION
TRANSPORTATION Air
TRANSPORTATION Land
TRANSPORTATION Marine

UTILITIES

VEGETATION
VEGETATION Floristic
VEGETATION Structural

WASTE
WASTE Liquid
WASTE Solid
WASTE Toxic
WASTE Sewage

WASTE Greenhouse gas
WASTE Heat

WATER
WATER Groundwater
WATER Hydrology
WATER Hydrochemistry
WATER Lakes
WATER Rivers
WATER Salinity
WATER Supply
WATER Surface
WATER Quality
WATER Wetlands

Qualifier words that may be added to any of the above:

Biodiversity
Classification
Conservation
Distribution
Exploration
Indicators
Inventory
Management
Mapping
Maps
Models
Monitoring
Networks
Planning
Production
Reference
Reports
Research
Reserve
Resources
Statistics
Surveys
Sustainability

Format Rules: At least one search word must be used**.** The main search word should be capitalised.

Example:

[Beecroft Peninsula Vegetation Survey]

FLORA Exotic Classification
FLORA Native Classification
VEGETATION Monitoring
VEGETATION Floristic Inventory
VEGETATION Structural Inventory

Element: GEOGRAPHIC EXTENT NAME(S)

Definition of Element:

Use either this element or the following element Geographic Extent Polygon, but not both. This is the ordinary name of one or more pre-defined, known geographic objects that would reasonably show the extent of geographic coverage of the data. It is not intended to show the exact extent. Use this element or the following element, whichever is the easier. Multiple entries are possible.

Allowable Content:

One or more of the following:

Australia excluding external territories
Australia including external territories
State or Territory
External Territories
Name and/or Number 1:1 000 000 Map Sheet (including offshore)
Name and/or Number 1:250 000 Map Sheet (including offshore)
Name and/or Number 1:100 000 Map Sheet
Name and/or Number 1:50 000 Map Sheet
Name and/or Number 1:25 000 Map Sheet
Local Government Areas
Statistical Local Areas
Major Catchment Basins
Interim Biogeographic Regionalisation of Australia (IBRA) Version 4.0
Interim Marine and Coastal Regionalisation of Australia (IMCRA)
Australian Navigational Charts
Oceans and Seas regions
Marsden Grid Squares

Format Rules:

There must be at least one, and may be many, entries for this element. Each entry should be selected from the objects from the classifications set out above. If none of the objects is a suitable representation of geographic extent, then the next field (Geographic Extent Polygon) should be used.

In the case of map sheets, the name and/or number of the sheet should be stated first.

In the case of Local Government Areas and Statistical Local Areas, please use names as found in the *Australian Standard Geographical Classification, ASGC, Catalogue 1216.0*, published by the Australian Bureau of Statistics.

In the case of major catchment basins, please refer to Major Drainage Basins as in Map 5 of the *Review of Australia's Water Resources, Division of National Mapping, Department of Minerals and Energy 1975*. This report is now out of print, but a digital representation of the Major Drainage Basins is available from AUSLIG (Australian Surveying and Land Information Group).

An Interim Biogeographic Regionalisation of Australia: A Framework for setting priorities in the National System of Reserves Cooperative Program (IBRA*), Version 4.0* Edited by R Thackway and I D Creswell, Australian Nature Conservation Agency, Canberra, 1995.

Australian Navigational Charts published by the Hydrographic Service RAN, various series and scales including Port Charts (1:5000 to 1:25000), Approach Charts (1:25000 to 1:75000), Coastal Navigation Series 1:150000 and 1:300000, Small Scale Series 1:1000000 and IHO Series 1:1500000, 1:3000000 and 1:10000000.

Ocean and sea regions are defined in *Limits of Oceans and Seas* published by the International Hydrographic Bureau, Special Publication 23, Draft Fourth Edition 1986.

Marsden Grid Squares in 10^0, 5^0 and 1^0 created by the World Meteorological Organisation and adopted by the International Oceanographic Commission.

Example(s):

Example 1: [Plant specimen records from the National Herbarium of NSW]

Australia including external territories

Example 2: [Beecroft Peninsula Vegetation Survey]

Jervis Bay 9027 1:100 000 Map Sheet
Kiama 9028 1:100 000 Map Sheet

Element: GEOGRAPHIC EXTENT POLYGON(S)

Definition of Element:

Do not use this element if the previous element has yielded a pre-defined extent that reasonably describes the extent of the dataset.

Use a set (or sets) of at least four coordinate pairs (a closed polygon) to record the latitude and longitude in decimal degrees of the geographic extent of the dataset where there is not a suitable predefined geographic extent available in the previous element (Geographic Extent Name).

Description of polygons may be made in up to four decimal places where necessary. Do not attempt to describe complex extents of irregular polygons in great detail - a minimum bounding rectangle is adequate. Multiple entries are possible.

Allowable Content:

Any number of sets of at least four coordinate pairs, where each polygon is closed by the repeat of the first pair.

Format Rules:

Express Latitude and longitude in decimal degrees to up to four decimal places.

Field Type: Real number

Length of Field: Maximum 1000 character string per polygon.

Other comments:

The coordinate pairs are to record latitude and longitude in decimal degrees. Specification to four decimal places would allow definition of 0.0001 degree areas. At the 36th parallel 1 degree roughly equals 91 kilometres. Thus 0.01 degrees roughly equals 1 kilometre on the ground. In most cases the level of precision given by two decimal places should be adequate. However, given that some organisations may use the ANZLIC guidelines internally in their organisation, allowing for four decimal places should allow specification of coverage to approximately 10 metres to occur. Repeating values allowed.

Example(s):

Example 1: [Bathymetry of the Gulf of Carpentaria and the Arafura Sea, Edition 1

-3.0 130.0, -3.0 149.0, -18.0 149.0, -18.0 130.0, -3.0, 130.0

Example 2: [Metropolitan Road Centreline Network]

-31.4553 115.5811, -31.4618, 116.4157, -32.4672 115.5654, -32.4740 116.4092, -31.4553 115.5811

Element: BEGINNING DATE
Element: ENDING DATE

Definition of Elements:

The earliest and latest dates for the records in the data. These are not necessarily the dates on which records were entered into a database.

It is possible that data may be recorded, for instance, many years after the collecting events – e.g. collections from Cook's Voyage to Australia in 1770. The dates that the events themselves occurred are those which are required in this element.

It is unnecessary to record the date of conversion of data from one form to another or from one database to another, or the date of conversion of the data to a specific format or system.

The intention of the elements are to record the age of the data themselves.

Allowable Content:

A date expressed as ddmmmyyyy, or the words "Not Known".

Where the year and month are known, but the day is not known, use mmmyyyy.

Where the year is known but the day and month are not known, just give the year.

Example(s):

Example 1:	30JAN1770
Example 2:	1889
Example 3:	Not Known

Element: PROGRESS

Definition of Element:

The status of the process of creation of the dataset.

If the data collection has concluded, pick the option "Complete". If the data collection is underway at the time of creation of the metadata record, use the option "In Progress". If the data collection has been proposed but has not commenced, use the option "Planned".

If the status of the process of creation is unknown, use the term "Not Known", but use this sparingly.

Allowable Content:

One of the following:

Complete
In Progress
Planned
Not known

Element: MAINTENANCE AND UPDATE FREQUENCY

Definition of Element:

This element is intended to describe the frequency of changes or additions that are made to the data after its initial completion.

If a dataset is maintained both daily and monthly, the most frequent rate of changes, i.e. daily, should be recorded. The changes to the data should be recorded, not the frequency of re-issue or publication.

Allowable Content:

One of the following:

Continual
Daily
Weekly
Monthly
Quarterly
Bi-annually
Annually
As required
Irregular
Not Planned
Not Known

Element: ACCESS CONSTRAINT

Definition of Element:

This element is intended to show any restrictions or legal prerequisites that may apply to use of the data, for example, entering into a licence/ royalty agreement. If access is unrestricted, that too should be stated.

If the data are only available in one particular delivery mode, e.g. available only through inter library loan, this should be indicated.

If payment is required, this should be stated.

If only some users are permitted access to the data, this too should be stated.

Example(s):

Example 1: [Plant specimen records from the National Herbarium of NSW]

Data are available for use in the Commonwealth Environment Portfolio. Permission has been received for release of the point data for Eucalypts on the World Wide Web. Endangered and vulnerable species, and species otherwise regarded as 'sensitive', must only be released as gridded information at a resolution no finer than 10' by 10'. The consultant shall notify Environment Australia of any taxa records regarded as 'sensitive'.

Example 2: [Beecroft Peninsula Vegetation Survey]

No restrictions.

Example 3: [Vegetation Survey data: Coongie Biological Survey, South Australia]

Data unrestricted. Should any publication arise out of the data the source should be acknowledged. SA Department of Housing & Urban Development require the opportunity to comment on any publication that uses the data - contact Custodian.

Element: LINEAGE

Definition of Element:

Lineage is a history of both the source data and the processing steps used to produce the data.

The source data used to produce the dataset may consist of one or more data sources. The history of the source data generally includes:

- a description of the source data
- the scale(s) of the source data
- the media type(s) of the source data
- the date(s) of the source data
- dates of various parts of the process.

The processing steps are the sequence of operational steps performed on the source data to arrive at the final dataset. The history of the processing steps generally includes:

- the data capture method(s)
- any intermediate processing method(s)
- the method(s) used to generate the final product
- any validation steps carried out on the data.

Allowable content:

Free text. If no answer is possible, use one of the following: Not Known, Not Documented or Not Relevant. The use of these, however, should be avoided wherever possible.

Field Type: Text

Length of Field: Maximum 2000 characters.

Example(s)

Example 1: [Plant specimen records from the National Herbarium of NSW]

Plant specimen records were received from the NSW National Herbarium comprising data from label information held by the Herbarium. These data cover a number of taxa and are obtained from time to time as required. Data are received and loaded into ORACLE Source tables. Validation for names is carried out against the Australian Plant Name Index.

Data then undergo a number of loading checks to check that the Taxonomic names exist and current, that the latitude and longitude are realistic values, and that the dates are realistic values and within range. Data are then loaded into the ERIS Specimen Module

Following loading, the data undergo a series of geographic validations. For those data that did not have a value for altitude, an altitude value was incorporated using a 1.5 minute DEM. Geographic and Altitude accuracy were estimated for each record from the data supplied. An off-shore test was conducted, and any records detected were flagged. Outliers in Latitude, Longitude and Altitude were detected using the ERIN validation tests. These use a modification of BIOCLIM to detect outliers in any of fifteen climate parameters. Any "suspect" records were identified and flagged. Any records flagged as "offshore" or "suspect" were returned to custodians for checking. Any corrections were made to the source tables and the data reloaded.

Data are then combined with data from a number of Australian State and Commonwealth institutions to form a composite, Australia-wide database.

Example 2: [Beecroft Peninsula Vegetation Survey]

The data were collected using 20 meter x 20 meter fixed quadrats. Total species counts, structure and other habitat data were also collected. The data were classified using Twinspan into groups comprising similar groups of species.

Element: POSITIONAL ACCURACY

Definition of Element:

Positional accuracy is an assessment of the closeness of the location of spatial objects in the dataset in relation to their true positions on the earth's surface.

The horizontal and vertical positional accuracy should be the assessed accuracy after all transformations have been carried out. This can be derived from a statistical analysis of tests e.g. root mean square error (RMSE) or standard deviation (SD).

A precise positional accuracy assessment may not always be possible. In these cases an intuitive estimate of the expected positional accuracy based on previous experience or expected likely maximum error is acceptable. In many cases this may be "just a feel" for the data but it is important to state this.

Positional accuracy may not be relevant to datasets that are indirectly geographically referenced. In this case choose Not Relevant from the list below.

Wherever possible the Geographic Datum used to determine the geocode position should be noted, eg. Aus 66, Aus 84, GDA 94, WGS 84.

Allowable content:

Free text. If no answer is possible, use one of the following Not Known, Not Documented or Not Relevant. These terms should, however, be used sparingly and information included wherever possible.

Length of Field: Maximum 2000 characters.

Example(s)

Example 1: [Plant specimen records from the National Herbarium of NSW]

Extremely Variable. All records include a figure (in meters) calculated for Geocode Accuracy and an Altitude Accuracy. Validation checks have been carried out in ERIN to look for outliers in geocoded position and altitude. Geocode determination unknown; geocode sources from collectors. Datum used also varies from record to record.

Example 2: [Forest Types]

Forest type boundaries were delineated on 1:15 000 or 1:25 000 aerial photos. Graphic (analogue) maps were hand digitised (0.5mm error) or scanned and adjusted where discrepancies compared to air photos occurred as for example misplacement of drainage, state forest roads, etc. The forest type boundaries were transferred from the aerial photographs to the digitised topographic maps to produce maps with the location of provisional forest type zones (an intermediate data set used in later processing). Sketch Master, Radial line plots, and Micromap 3D software are used in the processing. Questionable boundaries between forest types are checked in the field.

The forest type map was then compiled from this data set using a GIS package (ArcInfo).

Example 3: [Vegetation Survey data: Coongie Biological Survey, South Australia]

All sites have lat, long and AMG Coordinates (Datum: WGS94). Localities determined accurately using photopoints.

Element: ATTRIBUTE ACCURACY

Definition of Element:

Attribute accuracy is an assessment of the reliability of values assigned to features in the dataset in relation to their true 'real world' values. A list or summary of attributes of the dataset with a brief description of each should be included along with information on the accuracy of each attribute.
The attribute accuracy generally includes:

- the classification method used to assign values to features in the dataset
- an attribute accuracy assessment of how well the attributes conform to the classification method (often expressed as a percentage)
- an explanation of how the attribute accuracy assessment was determined.

A precise attribute accuracy assessment may not always be possible. In these cases an intuitive estimate of the expected attribute accuracy or the likely maximum error based on previous experience is acceptable. In many cases this may be "just a feel" for the data but it is important to state this.

Allowable content:

Free text. If no answer is possible, use one of the following: Not Known, Not Documented or Not Relevant. These terms should, however, be used sparingly and information included wherever possible.

Example(s)

Example 1: [Plant specimen records from the National Herbarium of NSW]

Variable. The taxonomy (including genus, species, infraspecies) has been checked by subject experts and should be largely correct. Other fields (collector name and number, date of collection, institution name and number) are as supplied by the custodian and should be largely correct

Example 2: [Forest Types]

The attribute of this data set is the forest type which is defined as any group of tree-dominated stands which possess a general similarity in composition and character. The classification in forest type was never intended as an ecological classification of forest vegetation in NSW. The forest type classification is based on economic and forest management considerations rather than on purely scientific ones. Considerations in using forest type classification were:

- the forest types are recognisable from aerial photographs
- each approved forest type is known to occur over appreciable land areas, and thus guarantees its representation at different scales for management planning as well as for day-to-day forest operations.
- each forest type is regarded as being sufficiently distinct from all others to warrant separate description.

There are 192 forest types identified and described by the Forestry Commission of NSW. Each type is given a distinctive number as a numerical reference.

These 192 types are divided into three major groups. Within each of these groups the types are further combined into assemblages of related types, called "leagues:. Whilst the leagues must be recognised as being essentially artificial groupings, they in most cases serve to unite those types which are most closely related to each other. The numbers given to forest types run ordinal through each league, with a gap in the series between most leagues so that, should further types be subsequently needed, these can be fitted into the existing framework of leagues and major groups.

The Aerial Photograph Interpreters do ground checks on forest types to check and correct identification. Their route is recorded on the aerial photographs together with the check points. The resource reports record where indicator species are present and any discrepancies that exist between the aerial photograph and the ground.

Element: LOGICAL CONSISTENCY

Definition of Element:

Logical consistency is an assessment of how well the logical relationships between items in the dataset, or spatial objects in the dataset, are maintained.

Spatial objects can be points, lines or polygons within the data that are used to represent true 'real world' features. When recording spatial objects into a dataset a number of inconsistencies can occur. An assessment for logical consistency documents for these inconsistencies. Tests are generally in the form of the following questions:

- Are all points labelled?
- Do lines intersect at nodes?
- Do lines cross unintentionally?
- Do all lines exist?
- Are lines duplicated?
- Do lines overshoot or undershoot?
- Are all lines labelled?
- Do all polygon boundaries close?
- Are all polygons labelled?
- Do any polygons have duplicate labels?
- Are all points, lines and polygons topologically related?

If the dataset is stored digitally then the tests for logical consistency can be carried out automatically using geographic information system software.

This element can also apply in the case of datasets where there are other logical relationships between items or objects (other than spatial objects) in the dataset. In such cases describe any tests carried on the relationships. Examples may be dates that occur in different fields - if the date given in one field says the project was carried out between years 'a' and 'b' but the date of recording of an attribute in another field is outside that range, then this is logically inconsistent; or records are out of range - if one field records the fact that data were collected in New South Wales, yet another field includes some records from Victoria, this is a logical inconsistency between the two fields.

Allowable content:

Free text. If no answer is possible, use one of the following Not Known, Not Documented or Not Relevant. These terms should, however, be used sparingly and information included wherever possible.

Example(s)

Example 1: [Forest Types]

The logical consistency tests done were:

- a test of valid values within each forest type
- a visual check of the maps, especially in the preparatory stages of map production, and
- a topological consistency check.

The valid value test checks for alien trees in well established tree formations.

The Photogrammetrist does a visual check to detect gaps in linework, to identify abnormal feature positions, correct line feature sharpness and to reposition displaced features as creeks out of their food plain or misplaced forest roads.

The GIS package (Arc/Info) was used to do topological consistency check to detect flaws in the spatial data structure and to flag them as errors. This check insures that all classified polygons are closed, nodes are formed at the intersection of lines and that there is only one label within each polygon, etc

Element: COMPLETENESS

Definition of Element:

Completeness is an assessment of the extent and range of the data with regard to completeness of coverage, completeness of classification and completeness of verification.

Completeness of coverage is an assessment of the proportion of the data available in its entirety (ie. spatial, temporal and attribute).

- Is the spatial data coverage complete for the entire dataset? If not what amount of spatial data is incomplete?
- Are attribute data available for the entire dataset? If not what amount of attribute data is incomplete?
- Does the data cover the whole time period? If not what is the temporal coverage of the data?

Completeness of classification is an assessment of how well the chosen classification method (refer to attribute accuracy) is able to represent the 'real world' features contained within the data.

- Is the adopted classification method exhaustive?
- Does the classification method generalise any features represented in the data? For example: Are there minimum area or minimum width rules used to represent features? (ie. roads less than 30 metres wide are represented as a single line); Must a lake be a certain area before it is included on a map at a scale of 1:100000?
- Are clusters of small areas amalgamated into one if they lie within a certain distance of each other?
- Are lines smoothed for presentation? If so, what method has been used?

Completeness of verification is an assessment of the amount of "work" (ie. field work or other methods) carried out to validate the correct representation of 'real world' features contained within the data.

- What is the extent and method of field verification carried out to validate both spatial and attribute data?
- Are the positions of any spatial objects in the data inferred? If so, what is the method of inference?

In completing entries for this element, use whatever concept of completeness as described above is most relevant to the data. If no answer is possible, use one of the following: Not Known, Not Documented or Not Relevant.

Example(s)

Example 1: [Plant specimen records from the National Herbarium of NSW]

Data largely complete for the taxa obtained, however not all taxa held by the herbarium have been obtained by Environment Australia. Data collection largely opportunistic, mainly from New South Wales, but includes some records from other States.

Example 2: [Forest Types]

Forest types are usually mapped at a scale of 1:25 000 following their delineation on aerial photographs of similar or larger scales (1:15 000). While quite small areas of special value can be readily identified in the field for individual management, areas less than 2 ha size, or 50 m in width were not represented.

Map legends are compact and standardised. Forest type numbers are shown on maps, but legends carry only limited descriptive information, such as actual major species occurring within types. Users of all maps are urged to consult map legends because of the ongoing changes to the classification (type numbers and description for types which have been recognised since the first edition of the Baur classification report (Research Note No. 71, revised 1989)). Maps without update classification will remain in use until time permits their revision.

Element: CONTACT ORGANISATION

Definition of Element:

This is the ordinary name of the organisation with which contact should be made to obtain the dataset itself, or to obtain more detailed information about the dataset. The contact organisation need not be the same organisation as the Custodian Organisation.

Element: CONTACT POSITION

Definition of Element:

The position title given by the Contact Organisation to the holder of the position who is required to answer questions about the dataset.

Allowable Content:

1. The ordinary name of the contact organisation in full.
2. The title of the position is preferred to the name of a person, as position titles tend to change less frequently than the occupants of positions.

If the contact organisation is also known by a common acronym, please include this at the end of the full name in round brackets.

If the contact organisation is usually known only by an acronym, use only the acronym without brackets.

Format Rules: Ordinary name of organisation followed where relevant by its acronym in round brackets**.**

Other comments: Multiple contacts may be allowed for the one dataset. Repeating values possible.

Example(s):

Example 1:

[Contact organisation]	State Herbarium of New South Wales
[Contact Position]	Database Manager
[Mail Address]	Mrs Macquaries Rd
[Suburb or Place or Locality]	Sydney
[State or Locality 2][Postcode]	NSW 2000
[Telephone]	02 9231 8111
[Facsimile]	02 9251 7231

Element: ADDITIONAL METADATA

Definition of Element:

An indication of where additional metadata about the dataset may be accessible. This could include a world wide web site address, a reference to another directory for example a theme directory, or a suggestion that more information should be sought from the contact position.

Example(s):

Example 1: See the “Blue Pages” Marine Directory at
<http://www.environment.gov.au/ cgi-bin/mcdd>