

Aristotle User Guide: How to record a controlled vocabulary

Key Terms: **Controlled Vocabulary, Classification, Reference Data**

User story: **I have a series of terms used to classify data that I want to record in my metadata registry, so I can inform people of what data I have and what it contains.**

My controlled vocabulary uses codes and categories to classify data. These codes should be able to be used for validating data, so I need to be able to access codes from an API. I need to be able to map between different codes when they record the same meaning.

If the codes in my vocabulary have changed over time, I need to record when different codes were valid for use in data. If my vocabulary is endorsed, I want to record who and when this controlled vocabulary was approved for use.

What metadata objects should I use?

Classifications, Conceptual Domains and Value Domains

Controlled Vocabularies are recorded in the Aristotle Metadata Registry using the **Classification**, **Conceptual Domain** and **Value Domain** metadata types. These metadata types can be linked to other metadata within the registry to make others aware of how data in the registry is coded or classified.

- A **Value Domain** is description of the possible values data can take and can include an ordered collection of codes with descriptive meanings.
- A **Conceptual Domain** is a description of how data is categorised and can include an ordered collection of named categories that describes how data is categorised.
- A **Classification** is an ordered collection of codes and categories for classifying data based on a particular topic. Classifications can be hierarchical to store complex categorisations of data.

One way to understand why controlled vocabularies and classifications are useful is to consider how why people classify things, such as clothing or books in a department store:

When stocking products in a store, department stores will group similar items together to help shoppers quickly find the products they want – for example, products could be grouped as cookware, toys, or clothes.

A child's cooking apron could be grouped in either one of these categories within a store, however shoppers will expect to be able to find the same product in the same place no matter which store they visit. So the company will set a strict classification for products so that users can easily find the products they want to buy.

How products are classified may also change based on how the clothing is processed. When on display, goods will be grouped by shoppers needs, when in a warehouse products may be grouped by manufacturer, and when products are shipped they may be grouped by weight.

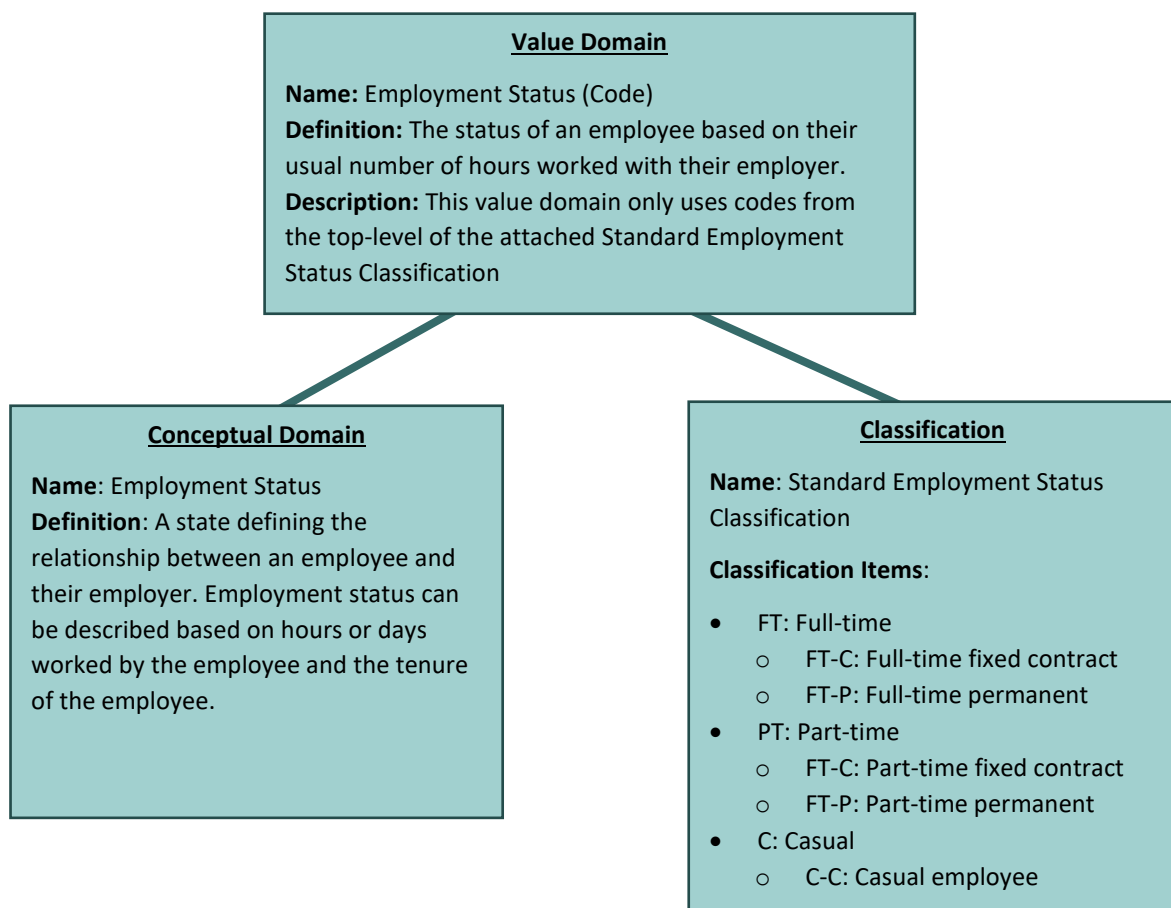
Having these classifications recorded and centrally managed means employees can easily find the information to do their job.

How Value Domains, Conceptual Domains and Classifications relate
Within the Aristotle Metadata Registry, a **Classification** is way of defining how your data is grouped together, including the standard codes and categories used to describe data. A Classification may record groupings for product data, employment data or health data and can be used to describe the appropriate use of these categories.

A **Value Domain** describes which values, codes and categories are used when recording data. When data is recorded using a fixed list of codes, these are recorded in a **Value Domain** as **Permissible Values**. When these values are reused at different levels in many different **Value Domains**, they can be stored as items in a Classification

Different business areas may use different codes when recording their data, due to different data practises or recording data in different language. By storing the basic definition of codes used in data in a **Value Domain**, users of the data will be able to continue to understand it.

A **Conceptual Domain** records the basic ideas used when categorising data for use. A **Conceptual Domain** can be purely descriptive, or it can also include the different lists of categories used when breaking down or aggregating data.



Examples of Controlled Vocabularies and Reference Data in Aristotle
 Below are some examples of how controlled vocabularies and reference data can be registered in the Aristotle Metadata Registry. Follow any of the links in the descriptions to see a complete example on the [Aristotle.Cloud](https://aristotle.cloud) registry.

Storing simple code lists using Value Domains

When data is recorded using coded values it is important that codes are recorded in a registry so that values can be interpreted during later analysis. In this example, numbers are used to record a persons' employment status in a company – with 1 meaning “employee”, 2 meaning “employer” and so on.

When this code list is attached to **Data Elements**, data in specific dataset can be properly interpreted.

Status in employment code N

Definition [ⓘ]
 A code set representing the person's position in relation to a job.

Representation

| Data Type | Number |
|--------------------------|--------|
| Format | N |
| Maximum character length | 1 |

Values

| | Value | Meaning | Start Date | End Date |
|----------------------|-------|-----------------------------------|------------|----------|
| Permissible Values | 1 | Employee | - | - |
| | 2 | Employer | - | - |
| | 3 | Own account worker | - | - |
| | 4 | Contributing family worker | - | - |
| Supplementary Values | 9 | Not stated/inadequately described | - | - |

Comments [ⓘ]

Guide for use:

Status in employment code N

Type: Value Domain

Identifiers: UUID - 67aabb98-158c-11e7-803e-0242ac110017
[meteor/270622/](#)

Origin URI: [Jump to original](#)

Stewardship Organisation: [Aristotle Demo Organisation](#)

Endorsed by:

- Community Services : Standard on 1st March 2005

[View registration history](#)

Last updated: 8th March 2015

Created:

Supersedes:

- None

Superseded by:

- None

Collections:

- None

View the example 'Status in employment code N' Value Domain online at:
<https://aristotle.cloud/item/270622/valuedomain/status-in-employment-code-n>

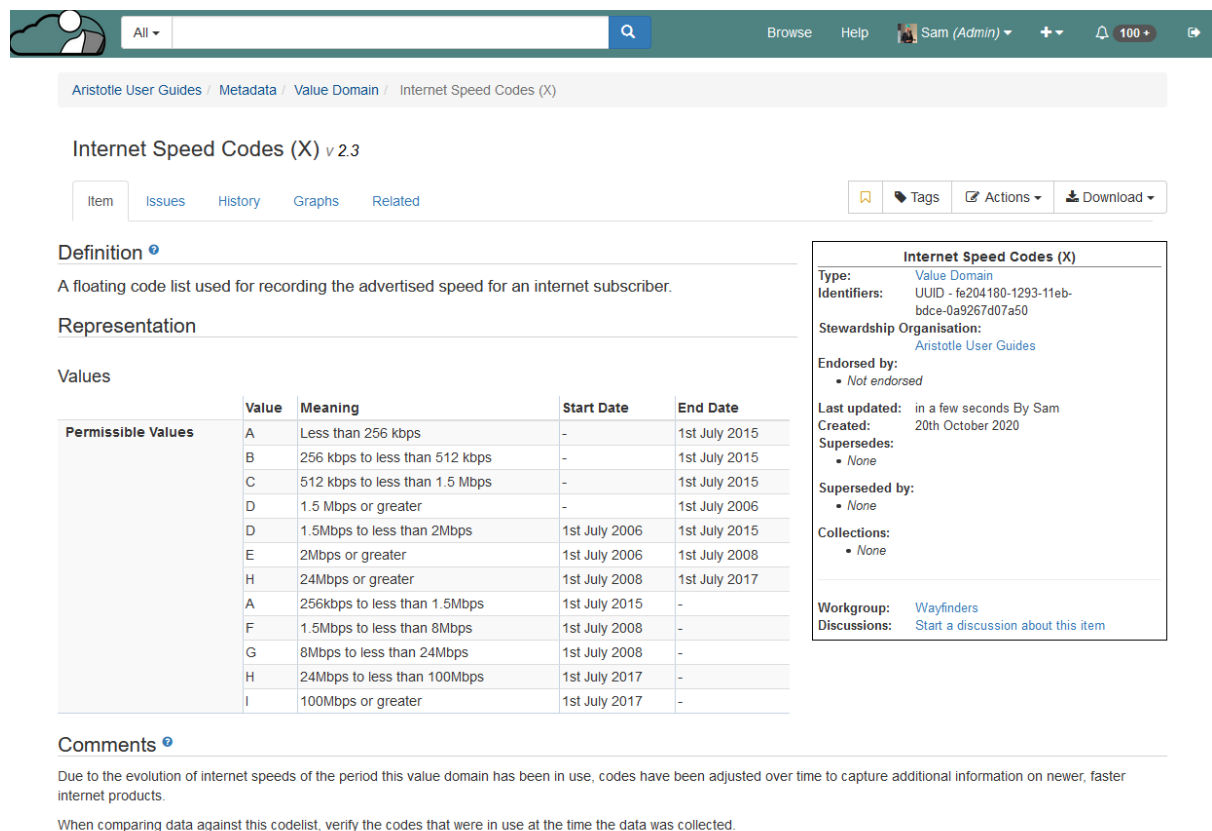
Storing code lists that change over time

As data can be collected over a range of years, metadata systems must be prepared to handle scenarios where data definitions and codes can change over time. This idea of tracking data as it changes over time is often called 'longitudinal data'.

Within the Aristotle Metadata Registry, code lists can be updated over time to ensure that users can accurately understand data as it was collected at different points in time.

In the example below, we have constructed a code list for recording Internet speeds provided to subscribers. Due to the evolving nature of the internet, internet speeds have increased dramatically over the past 10 years. Fortunately, the Aristotle Metadata Registry supports the recording of dates that codes were valid when managing code lists as Value Domains.

Because of this, definitions for Value Domains and Data Elements can be updated, without having to be recreated, which simplifies the management of 'longitudinal data' over time.



The screenshot shows the Aristotle Metadata Registry interface for the 'Internet Speed Codes (X)' Value Domain. The page includes a navigation bar with 'Aristotle User Guides / Metadata / Value Domain / Internet Speed Codes (X)'. Below the navigation bar, there are tabs for 'Item', 'Issues', 'History', 'Graphs', and 'Related'. A sidebar on the right contains icons for 'Tags', 'Actions', and 'Download'. The main content area is divided into sections: 'Definition', 'Representation', 'Values', and 'Comments'. The 'Values' section contains a table with columns for 'Value', 'Meaning', 'Start Date', and 'End Date'. The 'Comments' section contains a paragraph of text.

Internet Speed Codes (X) v 2.3

Definition ^o

A floating code list used for recording the advertised speed for an internet subscriber.

Representation

Values

| | Value | Meaning | Start Date | End Date |
|--------------------|-----------------------------|--------------------------------|---------------|---------------|
| Permissible Values | A | Less than 256 kbps | - | 1st July 2015 |
| | B | 256 kbps to less than 512 kbps | - | 1st July 2015 |
| | C | 512 kbps to less than 1.5 Mbps | - | 1st July 2015 |
| | D | 1.5 Mbps or greater | - | 1st July 2006 |
| | D | 1.5Mbps to less than 2Mbps | 1st July 2006 | 1st July 2015 |
| | E | 2Mbps or greater | 1st July 2006 | 1st July 2008 |
| | H | 24Mbps or greater | 1st July 2008 | 1st July 2017 |
| | A | 256kbps to less than 1.5Mbps | 1st July 2015 | - |
| | F | 1.5Mbps to less than 8Mbps | 1st July 2008 | - |
| | G | 8Mbps to less than 24Mbps | 1st July 2008 | - |
| H | 24Mbps to less than 100Mbps | 1st July 2017 | - | |
| I | 100Mbps or greater | 1st July 2017 | - | |

Comments ^o

Due to the evolution of internet speeds of the period this value domain has been in use, codes have been adjusted over time to capture additional information on newer, faster internet products.

When comparing data against this codelist, verify the codes that were in use at the time the data was collected.

Internet Speed Codes (X)

Type: Value Domain

Identifiers: UUID - fe204180-1293-11e5-bdce-0a9267d07a50

Stewardship Organisation: Aristotle User Guides

Endorsed by:
• Not endorsed

Last updated: in a few seconds By Sam

Created: 20th October 2020

Supersedes:
• None

Superseded by:
• None

Collections:
• None

Workgroup: Wayfinders

Discussions: Start a discussion about this item

View the example 'Internet Speed Codes' Value Domain online at:

<https://aristotle.cloud/item/605097/valuedomain/internet-speed-codes-x>

Storing concept lists using Conceptual Domains

Where code meanings have are complex or are reused in different code lists, they can be recorded separately in a **Conceptual Domain**. A Conceptual Domain allows users to record the different ways that data can be categorised, without specifying how these categories are recorded in data. This means that different data sets that use different codes, can still link to the same conceptual meanings, which makes comparison of different but related data easier.

This example, the Longitudinal Employment Study **Dataset**, records employment data collected over a ten-year period. In this **Dataset**, each **Distribution** records a data file for each year so that changes to data across time can be recorded.

This example also shows how the formats used to store the data have change over time, so users can understand how the data can be access and analysed.

Aristotle User Guides / Metadata / Conceptual Domain / Status in employment concepts

Status in employment concepts

Item History Graphs Related Download

Definition

A list of job categories based on their position and ownership of a company in relation to a job.

Value meanings

| Value | Meaning | Start Date | End Date |
|----------------------------|--|------------|----------|
| Employee | This code represents a person who works for a public or private employer and receives remuneration in wages, salary, or a retainer fee from their employer while working on a commission basis or for tips, piece-rates or payment in kind. An employee may also be a person who operates his or her own incorporated enterprise with or without hiring employees. | - | - |
| Employer | This code represents a person who operates his or her own unincorporated economic enterprise or engages independently in a profession or trade, and hires one or more employees. | - | - |
| Own account worker | This code represents a person who operates his or her own unincorporated economic enterprise or engages independently in a profession or trade, and hires no employees. | - | - |
| Contributing family worker | This code represents a person who works without pay in an economic enterprise operated by a relative. | - | - |

Status in employment concepts

Type: [Conceptual Domain](#)

Identifiers: [UUID - 02b068f0-0d13-11eb-ba99-0a66e1c05f44](#)

Stewardship Organisation: [Aristotle User Guides](#)

Endorsed by:
• [User Guide Standards Committee](#) : Recorded on 13th October 2020
[View registration history](#)

Last updated: a few seconds ago
Created: 13th October 2020

Supersedes:
• None

Superseded by:
• None

Collections:
• None

Related content

| Relation | Count |
|--|-------|
| Data Element Concepts implementing this Conceptual Domain | 0 |
| Value Domains Concepts implementing this Conceptual Domain | 1 |

View the example 2009-2019 Longitudinal Employment Study Data record online at:
<https://aristotle.cloud/item/605082/conceptualdomain/status-in-employment-concepts>

Using a Classification to store hierarchical codes and categories
A **Classification** can be used to record complex and deep data classifications.

This example **Classification** is based on an official classification of occupations – the Australia and New Zealand Standard Classification of Occupations. This classification stores over 1000 job categories at 5 levels

The screenshot shows the Aristotle Classification Viewer interface. At the top, there is a search bar and navigation links for 'Browse', 'Help', 'Toolbox', 'Log in', and 'Sign Up'. Below the search bar, the breadcrumb trail reads 'Aristotle User Guides / Metadata / Classification / Australian and New Zealand Standard Classification of Occupations'. The main heading is 'Australian and New Zealand Standard Classification of Occupations v 1.3'. There are tabs for 'Item', 'History', 'Graphs', 'Related', and 'Classification Viewer', with 'Classification Viewer' being the active tab. A 'Download' button is visible in the top right. The 'Classification Viewer' section has a search bar for terms. On the left, a tree view shows the hierarchy: 1 MANAGERS (with sub-items 11 Chief Executives, General Managers and Legislators, 12 Farmers and Farm Managers, 13 Specialist Managers, 14 Hospitality, Retail and Service Managers) and 2 PROFESSIONALS (with sub-item 21 Arts and Media Professionals, which includes 211 Arts Professionals, 2111 Actors, Dancers and Other Entertainers, 2112 Music Professionals, and 2113 Photographers). The main content area for '211 - Arts Professionals' includes a description: 'ARTS PROFESSIONALS communicate ideas, impressions and facts in a range of media to achieve particular effects, and interpret compositions such as a musical scores and scripts for performance.' It also lists an 'Indicative Skill Level' and provides educational requirements for Australia and New Zealand. Tasks include 'conceiving and developing ideas, styles, characters, emotions and themes and translating them through artistic presentation'.

View the Australia and New Zealand Standard Classification of Occupations online at:
<https://aristotle.cloud/item/605070/classificationviewer>

Because this classification is recorded as metadata within the registry, all **Value Domains** that link to this can be directly accessed through the registry as well.

Origin [®]

The original source for ANZSCO can be found online at: <https://www.abs.gov.au/ANZSCO><https://www.abs.gov.au/ANZSCO>

References

The Australian and New Zealand Standard Classification of Occupations (ANZSCO) and its components is Copyright © Commonwealth of Australia. Material here is reproduced for educational purposes, and is reproduced in compliance with the [Creative Commons 4.0 Licence](#).

Related content

| Relation | Count |
|---|-------|
| Value Domains referencing this Classification | 1 |

View metadata related to the Australia and New Zealand Standard Classification of Occupations online at: https://aristotle.cloud/item/605070/related/value_domains

References and Links

To ensure consistency between user stories across different organisations, the Aristotle Metadata Registry uses standard terminology when describing metadata types. Using standard terms makes it easier for people communicate by giving them a common vocabulary for describing how they manage their data and metadata.

The terms **Value Domain** and **Conceptual Domain** are defined in the internationally recognised ISO/IEC 11179 standard for metadata registries. The term **Classification** comes from the UN Statistics Division Classifications Model.

Links to these standards are included below.

ISO/IEC 11179: <http://metadata-standards.org/11179/>

UNSD Classification Model:

<https://unstats.un.org/unsd/classifications/expertgroup/egm2015/ac289-22.PDF>

When browsing or creating metadata within the Aristotle Metadata Registry, **Value Domains** and **Conceptual Domains** are included under “**Basic Metadata Registry Objects**” section.

| Basic Metadata Registry Objects | | | |
|--|------|--|---|
| These metadata types provide the core pieces for describing information recorded in a metadata registry. | | | |
| Create new | Help | Basic description | |
| Conceptual Domain | | A Conceptual Domain describes a set of ideas that can be recorded using codes when storing data. When linked to multiple Value Domains , a Conceptual Domain can be used to find similarities in different code sets. | <input type="button" value="+ Create"/> |
| Value Domain | | A Value Domain describes how to record the measurement of a particular type of data, either using a coded list of values or a description of the possible values. Value Domains can be linked to Data Elements that all share a common way of recording data, and its values can be linked to a Conceptual Domain to provide additional context. | <input type="button" value="+ Create"/> |

When browsing or creating metadata within the Aristotle Metadata Registry, **Classifications** are included under “**Classifications Management**” section.

| Classifications Management | | | |
|---|------|--|---|
| These metadata types provide classifications management support based on the Neuchatel and GSIM Classifications information models. | | | |
| Create new | Help | Basic description | |
| Classification | | A list of mutually exclusive categories representing values of the classification variable. | <input type="button" value="+ Create"/> |
| Correspondence Table | | A mapping of items between classifications to determine coverage and overlap of related Classifications. | <input type="button" value="+ Create"/> |

For more help on creating metadata in the Aristotle Metadata Registry, visit our Knowledge base: <https://desk.zoho.com.au/portal/aristotlecloudservicesaustralia/kb/articles/creating-metadata-items>

For more information on Aristotle.Cloud or the Aristotle Metadata Registry, visit us online at www.aristotlemetadata.com, contact the Aristotle team at hello@aristotlemetadata.com.



ARISTOTLE
metadata registry