ICT Project Guidance

Discovery:  
Information Resource   
Metadata, State, Context & Roles

Author:

Sky Sigal, Solution Architect

## Description

This document summarises a list of common states and roles that should be considered when analysing the capabilities of an information resource management system.

## Synopsis

Information Resources can be managed through a full information lifecycle with a small set of States and Roles with Permission to change them.

## Contents

[Description 1](#_Toc147309250)

[Synopsis 1](#_Toc147309251)

[Contents 2](#_Toc147309252)

[Introduction 4](#_Toc147309253)

[Draft 5](#_Toc147309254)

[For Review 5](#_Toc147309255)

[For Approval 5](#_Toc147309256)

[Rejected 5](#_Toc147309257)

[Approved 5](#_Toc147309258)

[For Categorisation 5](#_Toc147309259)

[For Release 5](#_Toc147309260)

[Released 6](#_Toc147309261)

[Replaced 6](#_Toc147309262)

[Merged 6](#_Toc147309263)

[Retired 6](#_Toc147309264)

[Archived 6](#_Toc147309265)

[Removed 6](#_Toc147309266)

[Restored 6](#_Toc147309267)

[Creator 7](#_Toc147309268)

[Owner 7](#_Toc147309269)

[Collaborator 7](#_Toc147309270)

[Reviewer 7](#_Toc147309271)

[Approver 7](#_Toc147309272)

[Endorser 8](#_Toc147309273)

[Maintainer 8](#_Toc147309274)

[Calibrator 8](#_Toc147309275)

[Member/Consumer 8](#_Toc147309276)

[Classification 9](#_Toc147309277)

[Metadata 9](#_Toc147309278)

[Media Type 9](#_Toc147309279)

[Technical Media Size 9](#_Toc147309280)

[Media Sizing 9](#_Toc147309281)

[State 9](#_Toc147309282)

[Category(s) 9](#_Toc147309283)

[Tags 10](#_Toc147309284)

[Copyright 10](#_Toc147309285)

[Calibration 10](#_Toc147309286)

[Linkage 10](#_Toc147309287)

[Sequence 11](#_Toc147309288)

[Similarity 11](#_Toc147309289)

[Hierarchy 11](#_Toc147309290)

[Ad-hoc Relationships 11](#_Toc147309291)

[Consent 11](#_Toc147309292)

[Appendices 12](#_Toc147309293)

[Appendix A - Document Information 12](#_Toc147309294)

[Images 12](#_Toc147309295)

[Tables 12](#_Toc147309296)

[References 12](#_Toc147309297)

[Review Distribution 12](#_Toc147309298)

[Audience 12](#_Toc147309299)

[Structure 12](#_Toc147309300)

[Diagrams 12](#_Toc147309301)

[Terms 13](#_Toc147309302)

## Introduction

Users use information systems to information.

They expect the information to be easily findable, preferably by multiple methods, in a forgiving way (e.g., forgives spelling mistakes).

They expect the information found to be valuable by being appropriate, complete, and current[[1]](#footnote-2).

For information within a system to be discoverable both the development process and organisation of the published information must be managed.

Risks

When Users cannot information, or find the information retrieved to be not appropriate, accurate, complete or current, they search elsewhere to find it.

The above outcome impacts the reputation of both the system and organisation.

Resolution

To publish high value resources a mature information lifecycle management process must be used, to permit multiple persons to contribute, review, approves the resources before they are published. Full lifecycle management systems permit the collection of end user feedback and commentary to inform iterative improvement to the resources.

To make published information resources more easily findable they require being categorised and contextualised using tags, categories and links.

To incrementally increase the value of information and keep it from devolving into becoming a reputational liability, a process to monitor and curate is required.

For the information state of information resources to be only changed by appropriately permitted users, Users are provided Resource specific Roles.

Information Resource States

Information Resources are managed through a full information lifecycle by multiple Roles being involved progressing them through a series of States.

The number of states is not fixed: the need for more or less States is determined based on the business use case.



Figure : Indicative Information Resource States to consider

A list of the most common States to consider include the following:

#### Draft

: a state enabling information resource Creators to invite other persons to have a Collaborator role associated to the resource to contribute to it.

#### For Review

: is not often used, but can be a formal state for a resource, such that it cuts off additional changes by Draft enabled Roles.

Either way, whether Draft or For Review is used, Persons can be invited as Reviewers to attach constructive comments to the resource.

#### For Approval

: for when Reviewer’s comments have been acted upon, and a User has been invited to be an Approver of the information resource.

#### Rejected

: for the Approver role to signal back to Collaborators who worked on the information resource that it is not ready to be published for consumption and returned to Draft state to continue working on it.

#### Approved

: for the Approver to signal to a Maintainer that it is ready to be categorised.

#### For Categorisation

: not often used, it can be a pre-step requiring a user to provide metadata for the resource, so that it is made more easily discoverable by being tagged, categorised, and in some specific cases, calibrated.

#### For Release

: the state to indicate that the resource is correctly categorised and ready for publication.

Note:

information Resources can be configured to have this state and be released in the future (e.g.: when a law comes into effect, or at the start of a term, etc.).

#### Released

: the initial state of a resource when categorised and released for consumption.

#### Replaced

: a previously published resource that has since been replaced with a later Released version. Resources in this state are displayed with links to the next version and latest version (while they initially are the same, this may not always remain the case).

Note:  
it is poor design to physically replace the previous copy such that it is no longer accessible if so desired.

#### Merged

: one of a set of two or more previously published resources that has been merged into a new Released information resource.

#### Retired

: an information resource that has not been Replaced or Merged with a later version and is considered unfit for continued use.

While remaining accessible, the information resource is displayed with information indicating it is no longer to be used.

#### Archived

: permitted system users may flag document for logical archiving.

This removes the resource from general user searching.

Note:  
Current best practice is to not remove information from systems (digital storage has become cheap), if privacy laws and concerns can be successfully met by removing them from discoverability.

#### Removed

: an information resource that has been removed from user access.

Note:

While it remains a use case-based design decision, it is recommended to also Remove all previous Replaced, Merged versions of the information resource.

#### Restored

: a permitted user may restore a Removed information resource back to the previous state.

Note:  
Implementation wise, it is recommended to consider making the Removed attribute separate from the Status attribute so that if later Restored, the document’s Published, Replaced, and Merged information is restorable.

Information Resource Roles

To change the state of the information resource, Users are invited to have Roles associated to the individual information source[[2]](#footnote-3).

Note:

The allocation of roles can be done by automation, based on default rules. For example, Users with specific Roles in a Group can be automatically assigned roles in information resources that belong to that group.

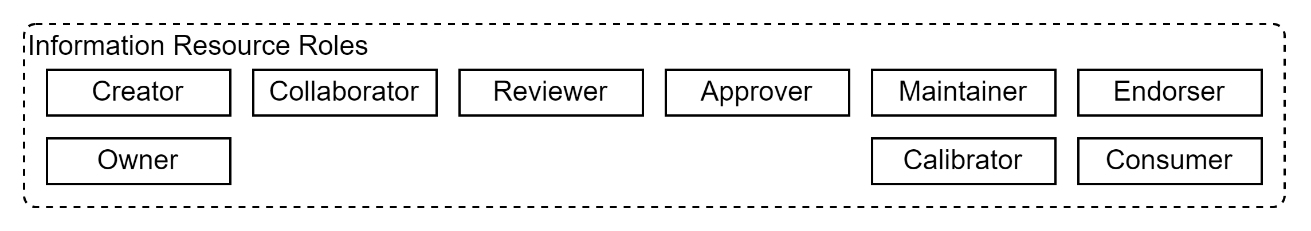


Figure : Example Information Resource specific Roles

#### Creator

: the first Collaborator on the development of an information resource in *Draft* state.

#### Owner

: a person accountable for the content of the information resource. It is common practice to automatically assign ownership to the Creator -- but can be assigned someone else later.

Note:  
Some systems permit multiple people being accountable.

#### Collaborator

: a User who is contributing to the development of the draft state information resource.

#### Reviewer

: a User with permission to Comment on and/or Contribute to the information resource.

#### Approver

: a User with the necessary permission to change the state of the resource to Approved or Rejected.

Note:  
The creator or any other collaborator can change an information resource’s state from Rejected back to Draft.

#### Endorser

: a User that can endorse, on behalf of a 3rd party authoritative body, an Approved information resource.

Note:  
Multiple endorsements can be given. For example, an information resource’s content may be approved by a cultural body, as well as a scientific body, etc.

#### Maintainer

: a User with the necessary permission to categorise the information resource by updating its metadata (tags, categories, links, etc.) to make it more easily discoverable when published.

A Maintainer can publish a resource, as well as replace it with a later version, or merge with other information resources into a new consolidated information resource.

A Maintainer can also Retire, Remove or Archive resources.

#### Calibrator

: not a common role, it is a role that can be allocated to a Subject Matter Expert (SME) so they can update specific metadata about an information resource. For example, in an assessment system, the Calibrator could rank a question to be suitable for a specific education level in a specific subject.

#### Member/Consumer

: a Person with permission to accesses the published Information Resources, provide feedback comments and ratings.

# Classification

A key aspect of making published information easier to find is by organising it, by the use of Metadata and establishing Links.

## Metadata

Information can be organised based on the attributes contained within a data record, or the attributes in separate record(s) (“metadata”) used to describe the information resource.

Note:  
The later approach offers more flexibility to find information resources efficiently, be linked in more place than one, etc.

### Categorisation

A traditional categorisation of Metadata purpose is as follows:

* descriptive,
* administrative, and
* structural.

### Standards

The classification of information is not a new concept, having started before digital documents became available: standards exist.

#### Dublin Core Metadata Initiative (DCMI)

: used to describe both digital and physical resources, known for its interoperability and versatility when used in conjunction with other metadata schemas.

#### Metadata Encoding and Transmission Standards (METS)

: generally used in digital libraries. It encompasses descriptive, administrative and structural metadata.   
It is known for its ability to add structure to a digital object. However, it is not as interoperable as other schemas. XML Based.

#### Metadata Object Description Schema (MODS)

: merges MARC bibliographic standards with Dublin Core's simplistic metadata terms. XML-based.

#### Text Encoding Initiative (TEI)

: primarily for the digital humanities field. XML based.

#### Visual Resources Association (VRA)

: for describing images. XML based.

Note:  
The most commonly referenced standard is DCMI.

### Common Attributes

Irrespective of schema standards, Metadata is commonly used to capture the following types of information:

#### Media Type

: information that describes whether the information resource is a Word document, text document, image, video, sound or other format.

#### Technical Media Size

: the size of the media, which informs on how long it may take to download a copy of the media.

#### Media Sizing

: the size of the media in media specific format (eg: number of pages) informing how long it may take to go through the material.

#### State

: as discussed earlier in this document.

#### Category(s)

: association to defined categories. Categories may be used to define Accessibility, and Usability qualities.   
A media resource may be categorised by language, and culture (mi-NZ).   
Categories for a Learning Resource may include pedagogical categories: *Curriculum, Subject, Level, Aspect, etc.*

Note: pedagogical categories may vary depending on other category values (for example, *Curriculum type*).

Note:  
mature systems offer the means of defining whether specific categories are enabled or not, as well as specifying from/to values. This permits categories to be enabled from specific dates (e.g.: to align with changes in applicable laws, scholastic term ends, etc. without requiring complex deployments of updated system and data on specific dates).

#### Tags

: association to ad-hoc defined terms of any kind (e.g.: “fun”, “triangles”).

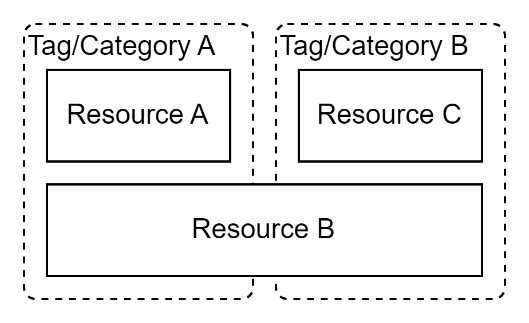


Figure : Organisation of information by Categories and/or Tags

#### Copyright

: anything that is published should be described as whether it can be reused or not.

**Important:**Copyright and copyleft are neither universal, nor capable of modelling accountable Acceptance of Obligations (e.g. Terms of Use) by an authoritative steward of information resource (e.g. an Iwi) for a specific use for a constrained duration.  
For example, the use of an image of an honoured warrior may not be copyright and instead may require *Consent* (see below) from a specific group (e.g. iwi) who may grant the use when limited to pedagogical reasons in a education provider context, for a specific agreed duration. A copyleft based permission may be unsuitable as it could imply the displaying of the image is provides fair use for it being printed on a t-shirt for sale to the public.

## Calibration

A specific type of classification is Calibration.   
A key difference is that if calibrating information resources is a desired outcome, they should be able to be calibrated against multiple scales. For example, a learning information resource should be calibratable against multiple the progression through multiple curriculums.

## Linkage

Another form of organising information is by linking information resources together.

Links between items provide the means of connecting information resources across categories.

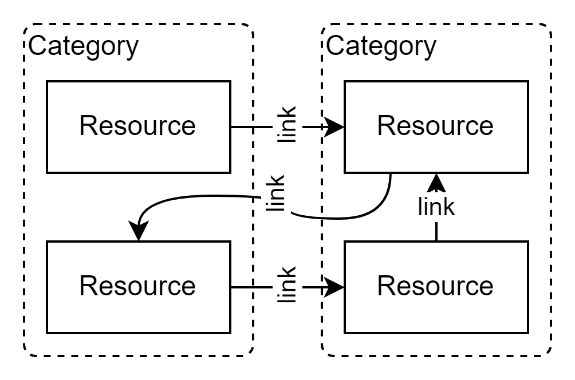


Figure : Categorisation by Links

The following are some types of links to consider:

#### Sequence

: e.g., Before/After

#### Similarity

: e.g., Synonym/Antonym

#### Hierarchy

: e.g., Parent/Child/Peer

#### Ad-hoc Relationships

: Recommended or Ad-hoc versions of Related to, of Interest to

#### Consent

: e.g. By/To/For (Beginning/Until).

**Important:**The ability for providing Consent is a means to address the limitations of blanket copyright/copyleft classifications, which do not provide a means to model and automate required acceptances of obligations, accountability, or durations of use.   
  
This is especially relevant to the inclusion and use of indigenous cultural resources (e.g., Māori and/or Pasifika media).

Appendices

Appendix A - Document Information

### Images

[Figure 1: Indicative Information Resource States to consider 5](#_Toc146779118)

[Figure 2: Example Information Resource specific Roles 7](#_Toc146779119)

[Figure 3: Organisation of information by Categories and/or Tags 9](#_Toc146779120)

[Figure 4: Categorisation by Links 10](#_Toc146779121)

### Tables

### References

**There are no sources in the current document.**

### Review Distribution

The document was distributed for review as below:

|  |  |
| --- | --- |
| Identity | Notes |
| Sandy Britain, Enterprise Architect |  |
| Amy Orr, Data Domain Architect |  |
| Jaleh Edwardson, Senior Business Analyst |  |
| Rodney Snell, Technical Lead |  |

### Audience

The document is technical in nature, but parts are expected to be read and/or validated by a non-technical audience.

### Structure

Where possible, the document structure is guided by either ISO-\* standards or best practice.

### Diagrams

Diagrams are developed for a wide audience. Unless specifically for a technical audience, where the use of industry standard diagram types (ArchiMate, UML, C4), is Appropriate, diagrams are developed as simple “box & line” monochrome diagrams.

### Terms

##### Attribute

: a data science term for single value within a digital record. Commonly referred to as a ‘field’.

##### Information Lifecycle

: the full lifespan of an information resource from conception to disposal, through the incremental stages of collaboration, review, approval, release, endorsement, replacement, removal or archiving.

##### Information Resource

: an element of information managed by a system for use by end users.

##### Service lifespan

: the full lifespan of a service from the moment it is first provisioning and released for use, through all its incremental releases to iteratively improve its capabilities, functionality and qualities, right through to its content being exported and transitioned for import into another system to meet a need for a continuity of service, and subsequent secure deprovisioning and decommissioning. It is a duration that can be measured in decades.

##### Subject Matter Expert (SME)

a specialist in a domain.

Refer to a project Glossary for further access to terms & acronyms.

Appendix B – Metadata Attributes

For consideration, the following attributes are sourced from the Dublin Core attributes.

While an interoperability standard, the list is expected to influence the design of the minimum set of system’s logical attributes that are required, if wanting to be able to be able to provide Dublin Core compatible attributes.

It is poor design practice to confuse the purpose of an *interoperability* standard with the purpose of an system’s *internal* information and data schemas. While one should adhere to an interoperability standard between a system and another, it is unacceptably limiting[[3]](#footnote-4) to also use the interoperability standard inside the system, as either a data persistence schema or logical element schema.

1. Contributor – "An entity responsible for making contributions to the resource".
2. Coverage – "The spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant".
3. Creator – "An entity primarily responsible for making the resource".
4. Date – "A point or period of time associated with an event in the lifecycle of the resource".
5. Description – "An account of the resource".
6. Format – "The file format, physical medium, or dimensions of the resource".
7. Identifier – "An unambiguous reference to the resource within a given context".
8. Language – "A language of the resource".
9. Publisher – "An entity responsible for making the resource available".
10. Relation – "A related resource".
11. Rights – "Information about rights held in and over the resource".
12. Source – "A related resource from which the described resource is derived".
13. Subject – "The topic of the resource".
14. Title – "A name given to the resource".
15. Type – "The nature or genre of the resource".

Note:  
 Dublin Core’s original attributes do not contain Structural attributes.

Appendix B – Copyright

By default, published work is protected by law[[4]](#footnote-5).

Copyright protects by law[[5]](#footnote-6), for an author’s life and an additional duration[[6]](#footnote-7), the following categories: literary, dramatic, musical and artistic works; the typographical layout of published editions; sound recordings; films; and communication works (such as TV/radio broadcasts and internet transmissions).

Exceptions exist for “fair use” towards research; private study; criticism or review; and reporting current events. Fair depends on the context. For example, in an education context, one can make a single copy, for lesson planning, and make copies of up to 3% or 3 pages, whichever is greater.

One can specify different terms of use. For example, Creative Commons[[7]](#footnote-8).

Important:  
When Copyright terms cannot be agreed upon, then requesting, considering, granting, and/or revoking Consent between parties is required.

1. Refer to ISO-25012 qualities of data to deliver. [↑](#footnote-ref-2)
2. Note that Information Resource Roles are specific to a Resource, and not to be confused with parent Group Roles. [↑](#footnote-ref-3)
3. Limits flexibility, maintainability qualities and design responsibilities. [↑](#footnote-ref-4)
4. [Understanding Copyright • Copyright Licensing New Zealand](https://www.copyright.co.nz/understanding-copyright) [↑](#footnote-ref-5)
5. In NZ, the Copyright Act 1994 [↑](#footnote-ref-6)
6. In NZ, 50 years. [↑](#footnote-ref-7)
7. [Homepage - Creative CommonsCreative Commons](https://creativecommons.org/) [↑](#footnote-ref-8)