

Public Record Office Victoria

GUIDELINE

REPRESENTING RECORDS IN VEOs

Version number: 1.0
Issue Date: 22/09/2023

Digital records which have permanent value must be transferred to Public Record Office Victoria (PROV) as VERS Encapsulated Objects (VEOs), at a time agreed between PROV and the public office.

This guideline is for technical developers building systems and tools to construct VEOs, and for agency record managers or system administrators creating VEOs for transfer to PROV. It describes how to construct VEOs so that they provide accurate and authentic representations of records from record systems. It outlines the criteria used by PROV when assessing VEOs for authenticity, a process known as archival assessment. Archival assessment is a component of the PROV VEO validation program.

This guideline supplements the technical requirements for VEO creation outlined in PROS 19/05 S4 Constructing VEOs and PROS 19/05 S5 Adding Metadata Packages to VEOs.

All references to VEOs are to Version 3 VEOs.

Table of Contents

Introduction	3
Public Record Office Victoria Standards	3
Applying this Guideline	3
1 Transforming Records into VEOs	4
1.1 Overview	4
1.2 Archival Assessment Criteria	5
2 About the Assessment Criteria	6
2.1 Identifying Records	6
2.2 Selecting Records for VEO Creation	7
2.3 Separating Records by Public Access Status	9
3 Organising Information in a VEO	10
3.1 Organisation of Information	10
3.2 Information Objects	11
3.3 Metadata Packages	12
3.4 Information Pieces in IOs	13
3.5 Content Files in IPs	13
3.6 IO Mapping to PROV ACM Item	13
4 Metadata	15
4.1 Satisfying the Mandatory Requirements of AGLS and AS-NZS 5478	15
4.2 AS-NZS 5478: the Preferred scheme	15
4.3 Implementing AS-NZS 5478	15
4.4 The PROV Minimum Metadata Specification (PROS 19/05 S2)	16
4.5 Inclusion of Additional Metadata Packages	16
5 The PROV Archival Control Model	17
5.1 PROV Archival Control Model	17
5.2 Agencies	18
5.3 Record Series	18
5.4 Items	19
5.5 Relationships between Entities in the PROV ACM	19

Introduction

Public Record Office Victoria Standards

Under section 12 of the *Public Records Act 1973*, the Keeper of Public Records ('the Keeper') is responsible for the establishment of Standards for the efficient management of public records and for assisting Victorian public offices to apply those Standards to records under their control.

Heads of public offices are responsible under section 13b of the *Public Records Act 1973* for carrying out a program of efficient management of public records. The program of records management needs to cover all records created by the public office, in all formats, media and systems across the organisation.

It is mandatory for all Victorian public offices to follow the principles and comply with the requirements of the Standards issued by the Keeper.

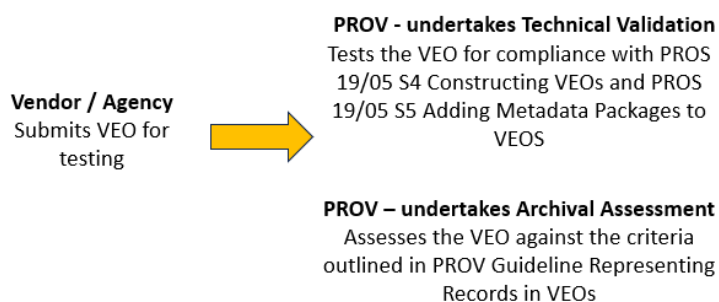
This guideline supplements the technical requirements for VEO creation outlined in PROS 19/05 S4 Constructing VEOs and PROS 19/05 S5 Adding Metadata Packages to VEOs.

Applying this Guideline

This guideline is intended for software developers building systems and tools to create Version 3 VERS Encapsulated Objects (VEOs), and agency record managers, system administrators or other staff responsible for creating VEOs for transfer to PROV.

It describes how to construct VEOs so that they provide accurate and authentic representation of records (their structure, metadata, and relationships) as they existed within a record system. It outlines the criteria used by PROV when assessing VEOs for accuracy and authenticity, a process known as archival assessment. It supplements the technical requirements for VEO creation outlined in PROS 19/05 S4 Constructing VEOs and PROS 19/05 S5 Adding Metadata Packages to VEOs. It is based on the recognition that VEOs may be technically valid but still fail to provide an authentic representation of a record.

Archival assessment is an essential element of the PROV VEO validation program:



Developers and agency staff responsible for VEO creation are encouraged to contact PROV for clarification or further discussion in relation to any of the matters covered in this guideline.

1 Transforming Records into VEOs

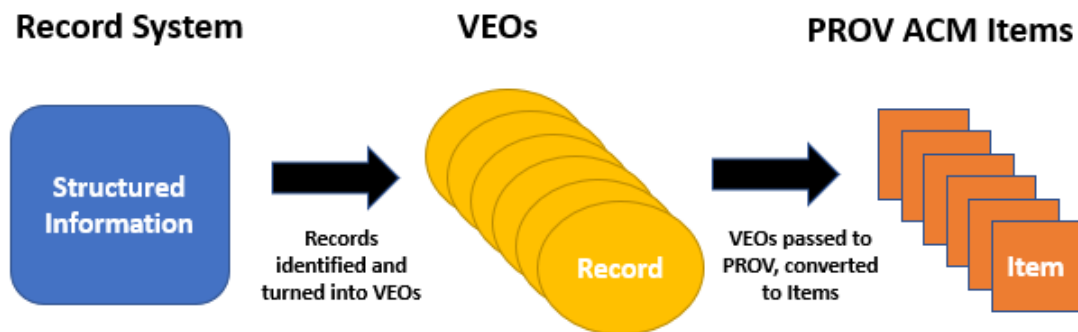
1.1 Overview

Record systems contain information. The task of constructing VEOs essentially involves identifying how that information has been structured into records and packaging those records as VEOs. The information within VEOs is extracted by PROV upon transfer and converted to Items within the PROV Archival Control Model (explained in Section 5 The PROV Archival Control Model). Items are accessed by agency and public researchers.

Generally:

1 Record = 1 VEO = one or more PROV Items.

The process can be conceptualised as follows:



For the purposes of these guidelines:

- A record can be thought of as information created, received, and maintained as evidence by an organisation or person, in pursuant of legal obligations or in the transaction of business. A record may comprise a single document or group of aggregated documents. 'Document' is used in the general sense to refer to any digital content e.g., a Word file, Excel spreadsheet, Image, PDF file, etc.
- 'Object' is used as a catch-all term to denote the different record types or entities that might be maintained within a record system. For example, containers and documents might both be referred to as objects.

This guideline provides guidance to ensure VEOs provide as authentic representation of records as possible.

Developers and agency staff responsible for VEO creation are strongly encouraged to contact PROV for clarification or further discussion in relation to any of the matters covered in these guidelines.

1.2 Archival Assessment Criteria

The criteria for the archival assessment component of VEO validation are summarised below.

Ref	Criteria
1	<p>The VEO provides an authentic representation of a record.</p> <p>This involves:</p> <ul style="list-style-type: none"> • The extent to which information was structured or organised as a record within a record system is represented in the VEO. • The extent to which relationships between records within the record system are captured in the VEO. • The extent and relevance of metadata capture. <p>The VEO must accurately represent the way information was structured or organised into a record by business users when the record was in active use. This means the VEO must include:</p> <ul style="list-style-type: none"> • Metadata seen by business users in the originating record system. • Metadata used to structure records within the record system and/or provide for discoverability (e.g., keywords or tags). • Where, a record consists of multiple documents, the arrangement of the documents and the relationships between them are reflected in the VEO.
2	<p>The process for selecting records for VEO creation should allow records to be selected based on how they were organised, grouped, or viewed by business users when in active use. Records can be selected for VEO creation based on classification term, keywords or tags associated with records, or some other user-specified criteria or rules.</p> <p>Records should NOT be identified and grouped for VEO creation by Disposal Status (Retention and Disposal Authority Class) alone.</p> <p>The grouping of Records for purposes of VEO creation needs to align with the PROV ACM Series entity.</p>
3	<p>Records can be selected for VEO creation according to the public access status that will apply to them post transfer to PROV.</p>

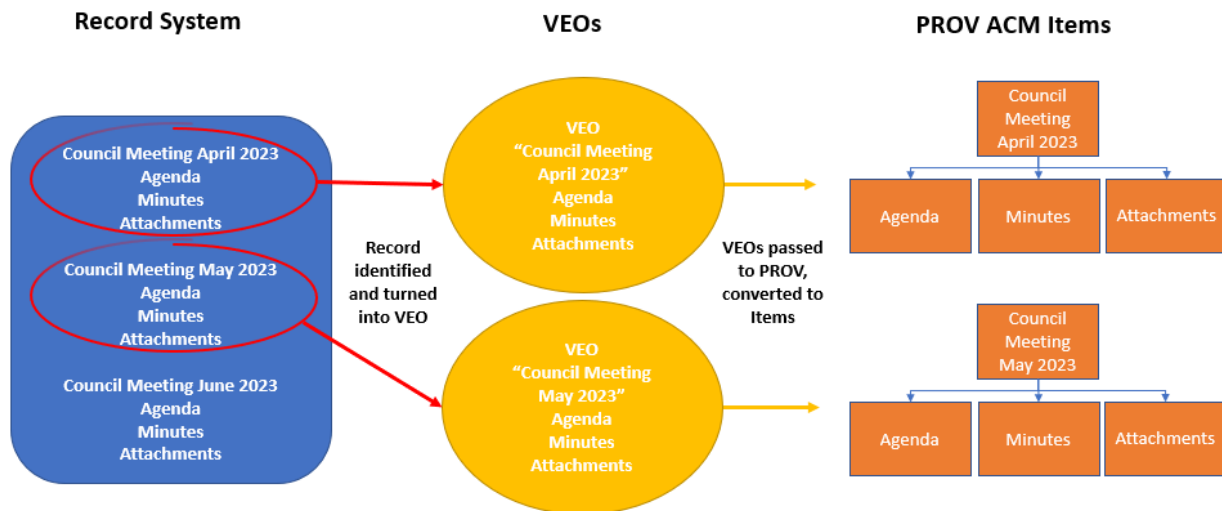
2 About the Assessment Criteria

2.1 Identifying Records

The record may be a single document or an aggregation of related documents that together provide evidence of decision making, a transaction, process, activity, or event. There is no single, fixed definition of what constitutes a record. What constitutes the record depends on the business activity being undertaken and how the records were created and managed by business users to provide evidence of decision making and actions.

For example, consider a record system that contained a set of documents relating to a sequence of Local Government Council meetings. It makes sense for each meeting to be a separate record. The contents of the record would be the documents relating to that individual meeting e.g., the agenda, minutes, and attachments.

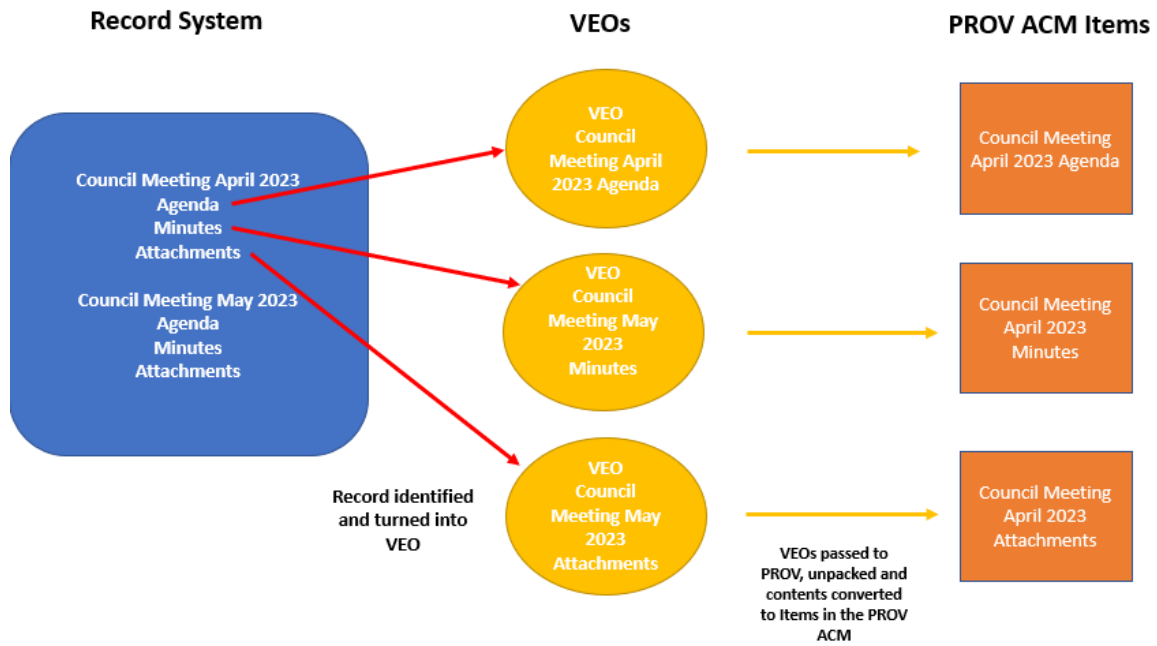
For example:



Document-centric approach

Alternatively, a document centric approach may be necessary if the VEO creation software cannot capture all related documents and the structural relationships between them, within a single VEO. In this scenario each document is captured as a single VEO, containing all the metadata, structural and relationship information to describe each document and its place within the agency recordkeeping system.

The previous example, could look like this:



Note: how each VEO must include all the metadata required to associate each document with its relevant Council meeting.

In the context of creating VEOs for transfer, business users are best placed to determine what constitutes the record, as they created and managed the records when in active use. PROV works in conjunction with business users to identify the record and this typically requires PROV being able to view the records within the original record system. This can be undertaken via an online screen share.

The organisation of information within a VEO is explained in more detail in Section 3 Organising Information in a VEO.

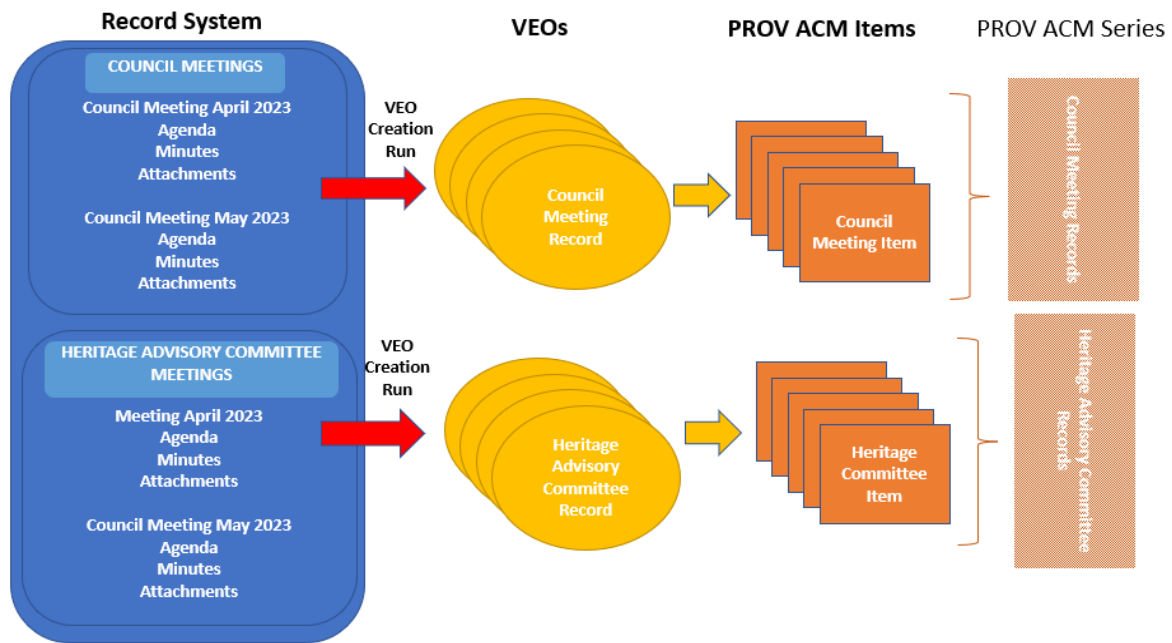
2.2 Selecting Records for VEO Creation

Record systems may provide structured ways of organising records, such as a formal business classification scheme, or views based on keywords or tags, either imposed by the business or user defined. It is important that the way records were organised or grouped by business users when in active use forms the basis for selecting or batching records for VEO creation.

It should be possible to select records for VEO creation based on classification term, keywords or tags associated with records, or other user-specified criteria, that determined how the records were organised, grouped, or viewed by business users when in active use.

Ideally, it should be possible to commence VEO creation from any point within a hierarchical structure, where a structure exists.

Preserving the way records were organised or grouped by business users for purposes of VEO creation can be conceptualised as follows:



Note: PROV preserves the original organisation or grouping of records once incorporated into the PROV ACM via the Series entity, depicted to the right in the above diagram (the Series entity is explained in detail in section 5.3 Record Series).

PROV recognises that multiple groupings or aggregations of records may be possible within a record system, and formal structures might not exist. Views may be dynamic, based on metadata keywords, for example, and vary from user to user.

Where a default or ‘master’ view exists then that is the arrangement to be replicated, if possible.

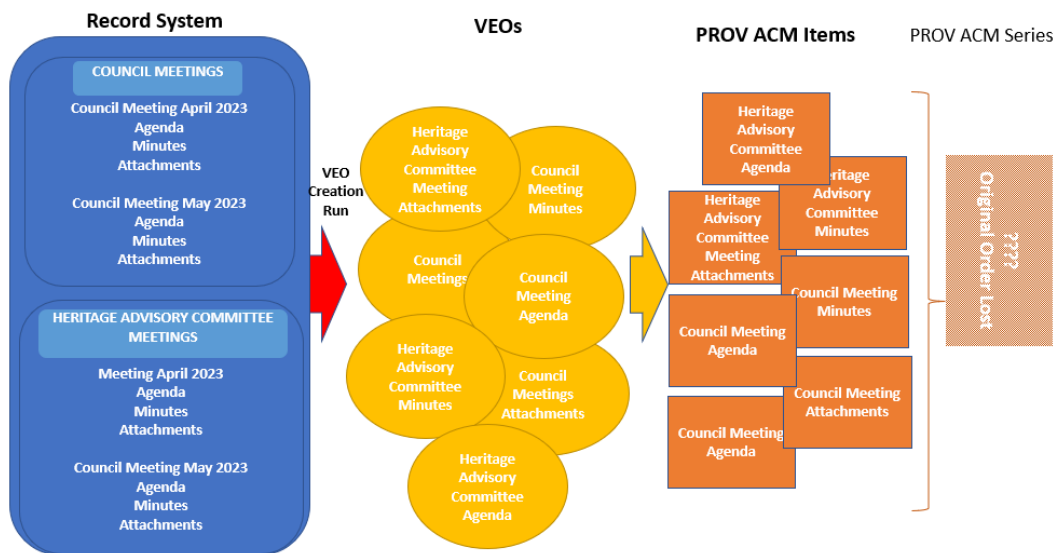
If such a view does not exist, then one can be selected for purposes of VEO creation.

PROV works in conjunction with business users to identify the basis on which records should be selected for VEO creation and the process can typically be undertaken via an online screen share.

2.2.1 Identifying Records by Retention Disposal Authority (RDA) class alone

Records should not be selected for VEO creation based on RDA class alone. Grouping records by RDA class doesn’t reflect the way records were grouped for purposes of conducting business. It also risks grouping unrelated records together.

The result can be illustrated as follows:



As shown, records relating to different business activities are mixed together – the original grouping or organisation of records by business users is lost, adversely impacting meaning and discoverability post transfer to PROV.

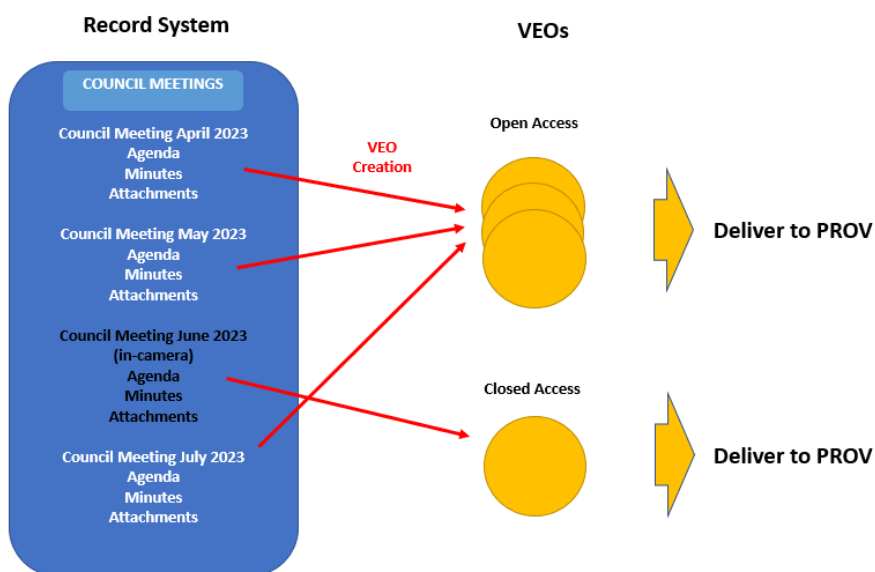
PROV recognises that some systems may be designed, out-of-the box, to identify records for VEO creation based on disposition status alone. Where a VEO creation process is based on RDA class alone, and cannot be varied, work will be required to re-organise the VEOs to reflect the way the source records were arranged when in active use.

2.3 Separating Records by Public Access Status

When selecting records for VEO creation, it may also be necessary to separate them according to the public access status that will be assigned to them post transfer to PROV.

For example, Council meetings may need to be divided into those that were public and hence open to the public post transfer to PROV, and those held in-camera, which will be closed to public access post transfer.

This can be illustrated as follows:



3 Organising Information in a VEO

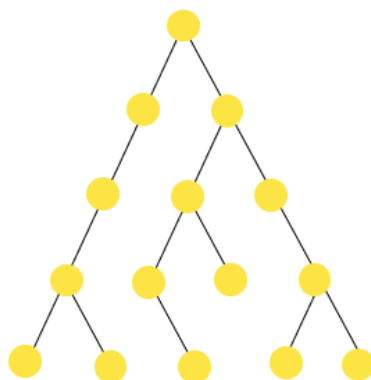
This section is intended to explain how information can be organised in a V3 VEO. It supplements PROS 19/05 S4 Constructing VEOs and PROS 10/05 S5 Adding Metadata to VEOs.

V3 VEOs are designed to describe, structure, and preserve digital information as it was managed by an agency, so it continues to be authentic, complete, and meaningful once exported from an agency system. The information within a VEO provides the source data for the creation of Items within the PROV Archival Control Model.

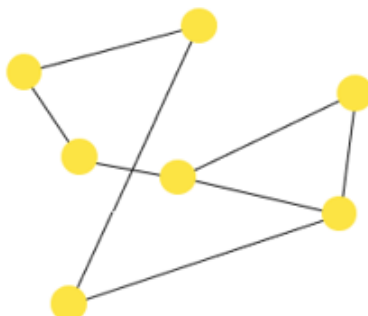
3.1 Organisation of Information

When constructing V3 VEOs, decisions need to be made about how to organise the information within a record system into individual records for packaging as VEOs. V3 VEOs are designed to be flexible about what is a record to maximise application across a range of scenarios (structuring the information is the function of the VEOContent.xml file - see PROS 19/05 S4).

In a V3 VEO, records can be organised into a tree to express their relationships and component parts. The creator of the record decides on the shape of the tree and what the nodes mean (the tree is physically represented in a V3 VEO in the VEOContent.xml file).

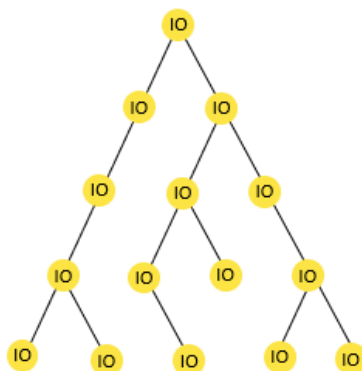


Note: You cannot depict a network (or graph) structure in a V3 VEO, that is:



3.2 Information Objects

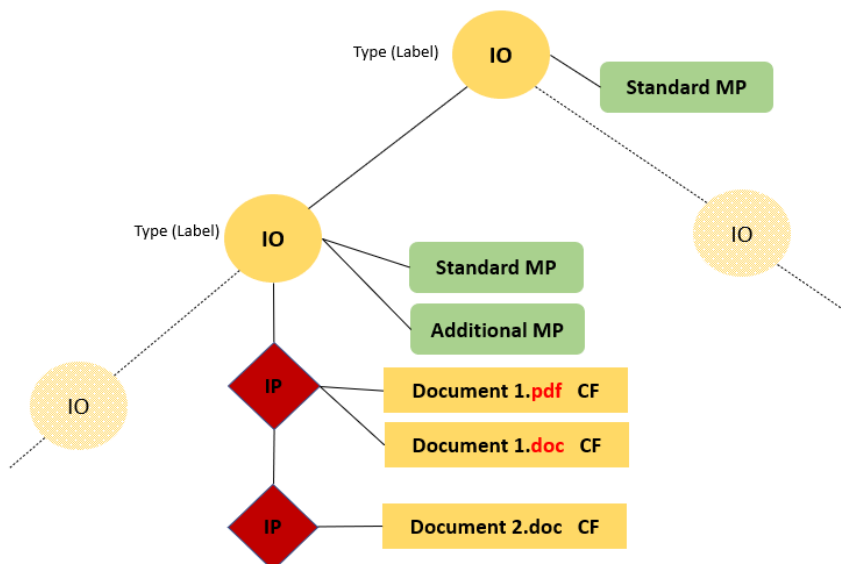
The nodes in the tree are called Information Objects (IOs). IOs are intended to be separate conceptual objects in the record e.g., containers in an EDRMS, records of business activities containing one or more documents, web pages, etc.



Information Objects have components:

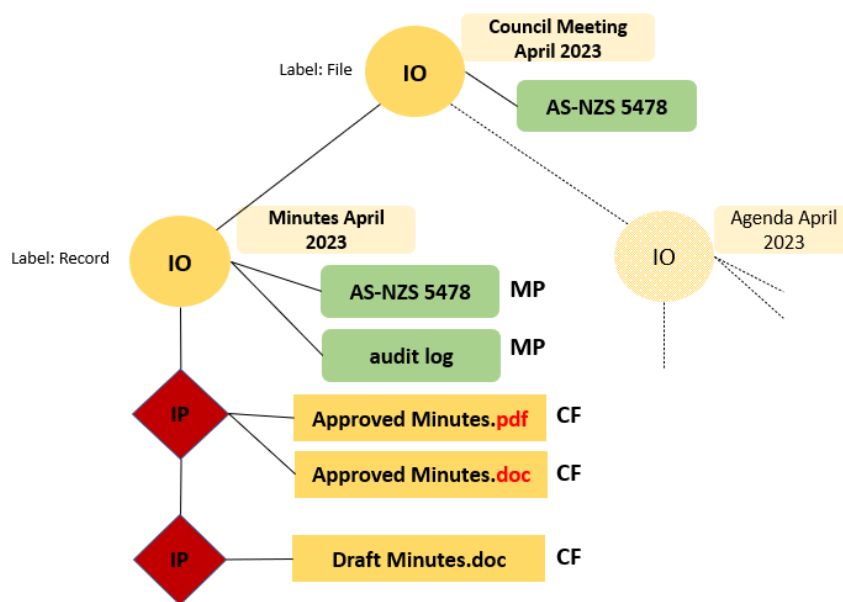
- A mandatory text label describing what the IO is. Each IO is labelled by a text string (called a type), to indicate the meaning of the node (e.g., 'Record', 'File'). Although it is called a type, it can really be any meaningful text string (e.g., a name) (see InformationObject-Type in PROS 19/05 S4 for further information).
- Metadata packages (MPs) containing metadata about the IO. It is mandatory to have at least one MP per IO (either AGLS or AS-NZS 5478 package).
- Information Pieces (IPs) representing the record content. IPs are optional, but if present each must contain at least one content file.
- Content Files (CFs), being the physical representation of a piece of record content.

For example:



- ➔ IOs are intended to be separate conceptual objects in the record (e.g., web pages, containers in an EDRMS)
- ➔ IPs are intended to represent separate components of an IO (e.g., the components of a web page)
- ➔ CFs are intended to represent different physical representations of the same file (e.g., Word/PDF).

In the previous Council meeting example, the VEO structure could look (in part) something like this:



- ➔ The root IO of each VEO is a container representing each Council meeting, containing the documents related to each meeting. There is a standard metadata package (AS-NZS 5478) and no additional metadata package.
- ➔ Beneath the root IO, the minutes and agenda are represented by individual IOs (the IO representing the attachments has been left out for simplicity).
- ➔ The IO representing the minutes has an additional metadata package containing audit information taken from the original system. It has two IPs, one containing the approved minutes in both pdf and doc format, and one containing the draft minutes.

3.3 Metadata Packages

Separating metadata into packages allows PROV to support multiple metadata standards.

- Each IO must contain at least one MP and may contain multiple MPs.
- Two standard metadata packages are defined: AGLS and AS-NZS 5478.¹
- Each IO must contain an AGLS or AS-NZS 5478 MP.
- Additional metadata packages can be included against any IO.

AGLS:

- Metadata standard primarily used to describe web resources.
- Derived from Dublin Core.
- Supported because it is a simple metadata standard and is likely familiar to Australian developers.
- Extended by PROV to include when a disposal decision should be made (year, action, mandate).

AS/NZS 5478:

- Australian / New Zealand Metadata Property Reference Set.
- Sophisticated record management metadata standard.
- Capable of expressing almost any recordkeeping concept required.

¹ See PROS 19/05 S5 Adding Metadata Packages to VEOs for further details re the metadata schemes and where to find them.

Additional metadata packages:

- Allow for IOs to be described using multiple metadata packages. For example, a photo could be generally described using AGLS, and then using specialised photography metadata.
- Allow unstandardised metadata from the source record system to be included (highly recommended by PROV).
- Will mean some metadata may be repeated, but this is not considered an issue by PROV.

3.4 Information Pieces in IOs

- IOs may contain multiple Information Pieces (IPs) representing record content (specific data such as a Word file or Image file).
- IPs are only used where an IO has record content.
- Multiple IPs can be used where the IO is represented by multiple pieces of record content.

3.5 Content Files in IPs

- Every IP contains at least one Content File (CF).
- A CF is the actual bitstream representing the Information Piece (e.g. Word files, JPEGs, etc).
- Multiple CFs are used in one IP when the same information is represented in different formats e.g. a Word file and a PDF file of the same document.

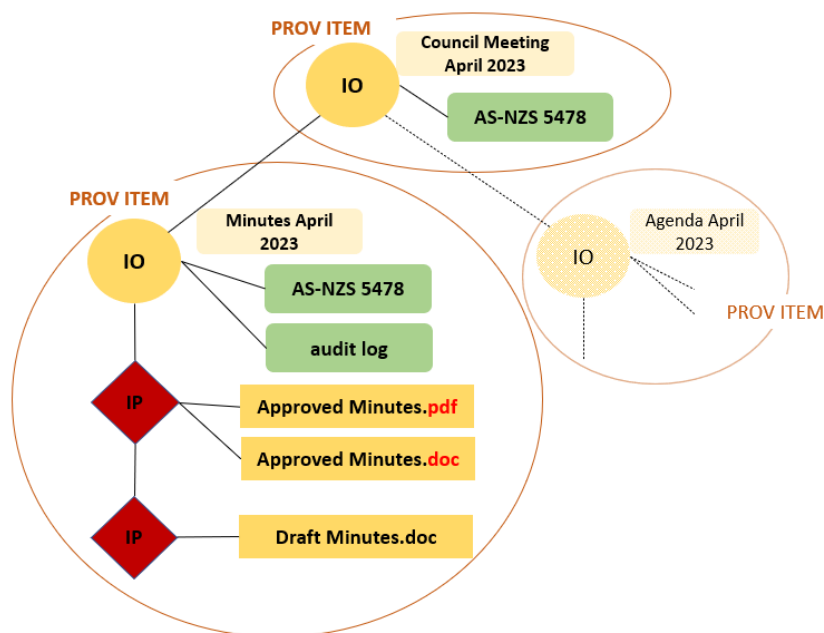
3.6 IO Mapping to PROV ACM Item

Information Objects map to PROV ACM Items:

- ➔ A VEO may contain one or more IOs.
- ➔ One IO = one PROV ACM Item.

This means a single V3 VEO may contain multiple PROV ACM Items.

In the Council meeting example, each IO maps to a PROV Item as follows:



The hierarchical structure of IOs within a VEO can be understood by the software that maps the IOs to the PROV ACM. The hierarchical arrangement of IOs within the VEO will be represented via the creation of parent/child relationships between Items in the PROV ACM.

The IOs in the above example would translate to 3 Items in the PROV ACM, related hierarchically, as follows:



- “Council Meeting April 2023” is a container Item, representing the grouping of all the documents related to that meeting.
- “Minutes April 2023” and “Agenda April 2023” are child Items, containing the minutes and agenda, respectively.
- “Minutes April 2023” contains the approved minutes (two instances, one pdf and one doc format) and the draft minutes.
- The standard metadata packages are mapped to the PROV ACM metadata scheme (which is based on AS-NZS 5478). Additional metadata packages are stored against the Item in the PROV ACM.

4 Metadata

4.1 Satisfying the Mandatory Requirements of AGLS and AS-NZS 5478

Each Information Object (IO) within a VEO must contain one of two standard metadata packages, either the Australian Government Locator Service (AGLS) metadata standard, or the Australian/New Zealand Recordkeeping metadata property reference set AS-NZS 5478.

Consult with PROV if any mandatory AGLS or AS-NZS 5478 metadata elements cannot be captured because the record system does not hold the necessary information.

4.2 AS-NZS 5478: the Preferred scheme

AS-NZS 5478 is the PROV-preferred metadata scheme because it provides a reference set of recordkeeping metadata intended to support records sustainability and systems interoperability.

AGLS, on the other hand, is designed to improve visibility and availability of *online* resources.

A freely available equivalent to AS-NZS 5478, the Australian Government Recordkeeping Metadata Standard (AGRkMS) and associated guidelines, is available from the [National Archives Australia website](#).

4.3 Implementing AS-NZS 5478

PROV requires AS-NZS 5478 implementations be flattened to the Record entity.² This means each IO within a VEO should represent a Record entity and each IO should contain one, and only one Record entity. This is to do with the way IOs map to the PROV ACM.³

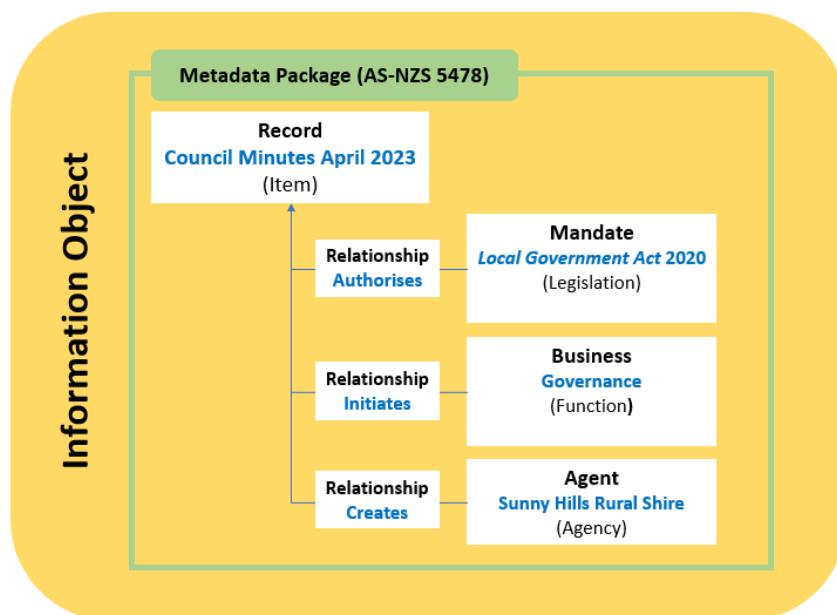
Agent, Mandate, Business and Relationship entities can still be included in a VEO, but they should be included within the Record entity IO, not as IOs in their own right.⁴

Here's an example of how a flattened implementation *could* look:

² This is also consistent with the Standard, which “allows for, but does not prescribe, a five-entity implementation Individual jurisdictions or organizations may choose to flatten ... metadata into a single entity, and that entity will always be the record entity” (AS-NZS 5478: 2015 p14).

³ An IO maps to one Item in the PROV ACM. Only Record entities map to the ACM. Hence, to map successfully, an IO should contain one, and only one Record entity. If an IO contained more than one Record entity, the Record entity with category Item has precedence for mapping to the ACM. If an IO contained more than one Record entity with category Item, only one of the entities would map to the PROV ACM. See Section 3.6 IO Mapping to PROV ACM Item for further information.

⁴ See PROS 19/05 S5 Adding Metadata to VEOs (Section 4 AS/NZS 5478 Metadata) for full details.



The Record entity category should be one of:

- Item, or
- File.

Record categories Archives, Archive, Series and Transaction Sequence are not required.

4.4 The PROV Minimum Metadata Specification (PROS 19/05 S2)

The PROV minimum metadata specification is just that, the bare minimum, and is intended for situations where available metadata is limited. Vendors and agencies are encouraged NOT to build to the minimum specification where additional metadata exists. PROV encourages the capture of rich metadata beyond the minimum specified.

4.5 Inclusion of Additional Metadata Packages

Any number of metadata packages, in addition to the standard packages can be included in a VEO. PROV strongly encourages the inclusion of additional metadata packages within the VEO. Examples include audit information and specialised metadata relating to the particular type of record such as image or geophysical metadata.

5 The PROV Archival Control Model

5.1 PROV Archival Control Model

PROV holds over 180 years of records - digital and physical, originating from thousands of record-keeping systems and government agencies. These records provide evidence of government administration and decision making. They must remain accessible to both government and public users indefinitely. It is therefore critical that different collections of records can be delineated from each other and the meaning of records – who created them and why – is preserved.

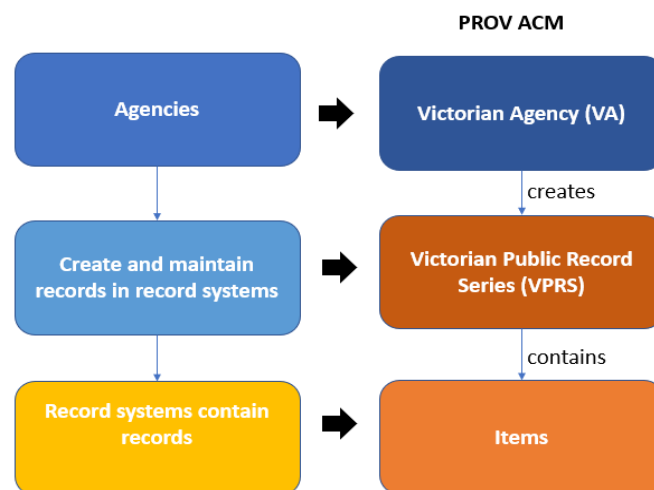
Researchers need to be able to see information about a particular topic, transaction, or subject, in the same way as the business user who created and managed the information when in active use. For example, they should see all the information ‘attached’ to a particular case or client file, just as the business user saw it. They need to be able to easily find relevant collections of information, and not get distracted by irrelevant documents.

- ➔ Digital records transferred to PROV are incorporated into a unified system of description, covering physical and digital records, from all manner of record systems and agencies, extending over 180 years. They must remain meaningful and discoverable.

PROV uses an archival control model (ACM) to achieve this aim. The ACM is the foundation for documenting information about records and their context, helping researchers to locate and understand government records, who created them and for what purpose. The ACM is a data model, metadata scheme and intellectual framework.

The PROV ACM is, like the systems of most archival institutions, based on the application of two related principles: The first, **provenance**, involves establishing the administrative context in which records were created and used. The provenance of records refers to their original creators and users and the records place within the record system maintained by their creators and users. The other principle, **original order**, relates to the relationships between records. It involves maintaining records in the arrangement in which they were when in active use.

Context and original order are documented within the PROV ACM via three key entities and the relationships between them: Agency, Series, and Item:



5.2 Agencies

Agencies are government organisations that create and manage records.

Agencies are identified and registered within the ACM by PROV. They are assigned a unique Victorian Agency (VA) number upon registration.

Note that when creating VEOs with AS-NZS 5478 metadata, the PROV ACM agency entity is not required to be represented in the VEO. That is, there is no need to include the PROV agency as an Agent entity (category Agency).

5.3 Record Series

Series are groupings of records, recorded or maintained by the same agency and which:

- Are in the same numerical, alphabetical, chronological, or other identifiable sequence; or
- Result from the same accumulation or recordkeeping process, perform the same function or have similar information content.

Series reflect how records were arranged when in active use within agency record systems. They reflect the groupings of records maintained by users when conducting business.

In the digital environment, each record system may contain one or more series. Series generally would not be split over multiple record systems. A record system will generally be considered to contain one Series if the system performs one function in an agency (e.g., a case management system manages cases). A record system will contain multiple series if the records in it relate to multiple functions or activities. Typically, this is when a record system is a general-purpose system that contains records relating to multiple functions/activities within an agency (e.g., an EDRMS or EDMS).

Records transferred to PROV need to be arranged (or grouped) according to Series. When creating VEOs, records are selected from a particular series in a record system for export as VEOs.

Note that the organisation of records into Series is NOT based on Retention and Disposal Authority classes. RDAs are a means of identifying which records are eligible for transfer to PROV, rather than a grouping mechanism reflecting active use (note too, that RDAs can change over time, they aren't fixed forever).

Series are identified by PROV in conjunction with agency business users. Series are identified and registered within the ACM by PROV. They are assigned a unique Victorian Public Record Series (VPRS) number upon registration.

Note that when creating VEOs with AS-NZS 5478 metadata, the PROV ACM Series entity is not required to be represented in the VEO. That is, there is no need to include the PROV Series as a Record entity (category Series).

Although Series entities are not required to be represented in VEOs, they can be (highly recommended they are) identified by their PROV registration (VPRS) number in the 'RDF about' element along with the Series consignment number (see below).⁵

Subgroupings within Series are required to separate records according to their public access, should there be a mix of open and closed records within the Series. The subgroupings are known as Consignments. The Consignment id is issued by PROV and should also be identified within the 'RDF about' attribute.

⁵ See PROS 19/05 S5 Adding Metadata Packages to VEOs section 2.6.1 "rdf:about attributes".

5.4 Items

Items represent records transferred to PROV from agency record systems. An Item may consist of one or more documents or other type of computer object, a container or file, or a combination of both. It depends on how the record has been defined within the originating record system.

Unlike Agencies and Series, Items are created from information extracted from VEOs. Information is extracted from record systems, used to construct VEOs, which are then processed at PROV to create Items in the ACM. As explained previously, Items map from Information Objects (IOs) within VEOs.

Item Metadata

PROV maintains descriptive, technical & administrative metadata about Items within its archival management system.

Descriptive metadata, including relationship metadata and any keywords, tags or location type metadata (indicating the location of the record within the originating record system) is the key for preserving the context in which a record was created, maintained, and used when in active use, and for enabling discoverability.

The key descriptive metadata immediately visible to agency and public users post transfer includes:

- Title
- Description
- Date(s)
- Identifier(s)

5.5 Relationships between Entities in the PROV ACM

Agency to Series Relationships

- Agency to Series creates / created by
- Agency to Series responsible for / responsibility of.

Relationships between Agencies and the Series that they created and maintained are established within the ACM by PROV.

Series to Item Relationships

- Series to Item contains / contained in.

The relationship between Series and Items within the PROV ACM is established by PROV at time of ingest to the PROV digital archive.

Item to Item Relationships

- Item to Item Contains / Contained in.

Hierarchical relationships between Items are established during VEO processing at PROV at the time of ingest to the digital archive. They are identified from the structure of the Information Objects (IOs) within a VEO. Currently, they can only be created between IOs within the same VEO but this could change in the future.

Lateral relationships between Items are not currently maintained within the PROV ACM but are considered a possible future enhancement.

PROV strongly encourages all relationships between records within a record system be included in the Information Object (IO) metadata within the VEO, regardless of whether they are currently supported in the PROV ACM.

Copyright Statement

© State of Victoria 2023



Except for any logos, emblems, and trademarks, this work is licensed under a Creative Commons Attribution 4.0 International license, to the extent that it is protected by copyright. Authorship of this work must be attributed to the Public Record Office Victoria. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/legalcode>

Disclaimer

The State of Victoria gives no warranty that the information in this version is correct or complete, error free or contains no omissions. The State of Victoria shall not be liable for any loss howsoever caused whether due to negligence or otherwise arising from the use of this Standard.

