Queensland Government Enterprise Architecture

Identification and classification of information assets guideline

Final

June 2016

V2.0.0

PUBLIC

Document details

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 1. **Security classification** | 1. PUBLIC | | | | | |
| 1. **Date of review of security classification** | 1. June 2016 | | | | | |
| 1. **Authority** | 1. Queensland Government Chief Information Officer | | | | | |
| 1. **Author** | 1. Queensland Government Chief Information Office | | | | | |
| 1. **Documentation status** |  | 1. Working draft |  | 1. Consultation release | 1. 🗹 | 1. Final version |

Contact for enquiries and proposed changes

All enquiries regarding this document should be directed in the first instance to:

Queensland Government Chief Information Office  
[qgcio@qgcio.qld.gov.au](mailto:qgcio@qgcio.qld.gov.au)

Acknowledgements

This version of the *Identification and classification of information assets guideline* was developed and updated by Queensland Government Chief Information Office.

Feedback was also received from a number of agencies, which was greatly appreciated.

Copyright

*Identification and classification of information assets guideline*

Copyright © The State of Queensland (Department of Science, Information Technology and Innovation) 2015

Licence

[](http://creativecommons.org/licenses/by/4.0/)

This work is licensed under a Creative Commons Attribution 4.0 International licence. To view the terms of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. For permissions beyond the scope of this licence, contact [qgcio@qgcio.qld.gov.au](mailto:qgcio@qgcio.qld.gov.au).

To attribute this material, cite the Queensland Government Chief Information Office.

The licence does not apply to any branding or images.

Information security

This document has been security classified using the Queensland Government Information Security Classification Framework (QGISCF) as PUBLIC and will be managed according to the requirements of the QGISCF.

Contents

[1 Introduction 4](#_Toc453241087)

[1.1 Purpose 4](#_Toc453241088)

[2 Purpose 4](#_Toc453241089)

[2.1 Audience 4](#_Toc453241090)

[2.2 Scope 4](#_Toc453241091)

[3 Background 4](#_Toc453241092)

[4 Definition of an information asset 5](#_Toc453241093)

[5 Guidelines and discussion 5](#_Toc453241094)

[6 Classification of an information asset 8](#_Toc453241095)

[6.1 Using the Information classification framework 8](#_Toc453241096)

[6.2 Classifying information assets 10](#_Toc453241097)

[7 Application specific considerations 12](#_Toc453241098)

[7.1 eDRMS 12](#_Toc453241099)

[7.2 Static website content and web content management systems 13](#_Toc453241100)

[8 Information asset identification and classification form 14](#_Toc453241101)

# Introduction

A Queensland Government Enterprise Architecture (QGEA) guideline provides information for Queensland Government agencies on the recommended practices for a given topic area. Guidelines are generally for information only and agencies are not required to comply. They are intended to help agencies understand the appropriate approach to addressing a particular issue or doing a particular task.

This guideline specifies how to correctly identify and classify an information asset. The last section contains a checklist to assist with the identification of information assets.

## Purpose

Identification of information assets is an important step defined within the Queensland Government ICT planning methodology.

## Audience

This document is primarily intended for:

* chief information officers (CIO)
* information management directors
* information owners
* information managers
* information asset custodians
* data managers
* information users.

## Scope

### In scope

This guideline relates to all of the domains within the information layer of the QGEA.

It applies to all Queensland Government departments.

### Out of scope

The following are out of scope of the current guideline:

* classification of information assets to schemes other than the QGEA Information classification framework
* classification of information assets for security purposes – this is covered in[Information Standard 18: Information security (IS18)](https://www.qgcio.qld.gov.au/products/qgea-documents/549-information-security/2704-information-security-is18policy).

# Background

The [Information asset custodianship policy (IS44)](https://www.qgcio.qld.gov.au/products/qgea-documents/548-information/2632-information-asset-custodianship-policy-is44) requires agencies to identify the agency’s holdings of information assets and establish an information asset register. It is the role of information asset custodian to ensure that information assets have been identified, classified and recorded in an information asset register.

The identification and classification of information assets assists with identifying the scope, types, use and functions of an agency’s information asset holdings. It helps to facilitate the access and reuse of information and reduces any unnecessary duplication of information assets.

In addition, as part of the process of identifying and collecting information assets, attention should also be paid to the quality and consistency of the information stored in the information asset register/s as it has a direct bearing on the level of usefulness to the organisation.

# Definition of an information asset

An information asset is defined within the QGEA as:

1. *An identifiable collection of data stored in any manner and recognised as having value for the purpose of enabling an agency to perform its business functions, thereby satisfying a recognised agency requirement.*
2. *Data or information that is referenced by an agency, but which is not intended to become a source of reference for multiple business functions is not considered to be an information asset of the agency...*

# Guidelines and discussion

To ensure that an information asset is identified and named at a consistent level of detail, it is recommended that agencies apply the following guidelines.

| 1. Guideline | 1. Discussion |
| --- | --- |
| 1. An information asset is a logical concept | 1. An information asset is a logical concept; not a physical manifestation. Although it is logical, it also has tangible business meaning. 2. To recognise the logical nature of an information asset, imagine the agency as a single entity, ignoring the underlying applications and technologies. In this ideal world, there would only be a limited number of assets that represent key business concepts. 3. For example, a training organisation deals with people; however, the organisation recognises that there are different types of people that are critical to their business. Analysis of their application reveals these types to include students, teachers, administrators, maintenance staff etc. 4. Common business nouns (e.g. students, teachers from the above example) can provide an indication of the scope and number of the agency’s information assets. |
| 1. An information asset should be named using nouns | 1. The name of an information asset should not contain a verb. Rather, the following format should be used: 2. Asset name = subject noun (typically singular) + format noun (inherently a collection, group or plural) 3. The **subject noun** is the main or primary subject of the information, using your agency’s own terminology. Avoid the use of nouns such as information or management if possible. ‘Information’ is too generic to provide a business context and ‘management’ in particular can be easily confused with a business process. 4. The **format noun** describes the system-independent structure and content type of the information asset. Avoid using terms like database, table, cube or system. The best names are those with the most relevance to your agency’s business. 5. Example of format nouns include: set, register, list, hierarchy, feedback, catalogue, inventory, ledger, content, dataset, index, report, tabulation, matrix, correspondence, library, imagery and map. Agencies will also have specific terms such as the term ‘cadastre’ used by the Department of Environment and Resource Management. 6. The following are appropriate examples: carers register, customer service feedback, complaints register, crime location map, registered office list, service delivery hierarchy, policy library etc. |
| 1. An information asset is named independently of any system or application | 1. In the majority of cases, a single application will contain multiple information assets. Information assets are considered to be conceptually separate from their implementation and exist regardless of the application that underpins it. 2. For example, the Department of Child Safety has the integrated client management system (ICMS). This system contains a number of information assets such as a child register and foster carer register. Note that these assets have been named independently of the system that underpins them. 3. In this example, it would not be acceptable to group the information assets as one asset called the ICMS database. |
| 1. An information asset should be named using the agency’s common business terminology | 1. Most agencies have unique business terminology. Information assets should be named in accordance with this terminology. 2. For example, if the business refers to ‘clients’ rather than ‘customers’ use client as the subject noun of the information asset under consideration. |
| 1. An information asset has value if it is actively used by a business or decision making process or exhibits a lifecycle | 1. An information asset has value for an agency where an item of information contained in the asset is:  * required as input to a business process or is the output of a business process * used to evaluate a rule or condition * subject to a typical information lifecycle (plan, create, store, access, use and maintain) where the agency is responsible for some or all stages of the lifecycle. |
| 1. Reference information sources are not considered to be an information asset of the agency | 1. Information will only become an information asset of the agency if it is held and maintained by the agency according to the previous guideline. External reference material or information provided for context is generally not considered an information asset. 2. For example, information about the percentage of the children across the population currently in foster care would be an information asset of the Department of Child Safety. In particular the count of children in care would be a critical item of information. However, the total population of Queensland is only referenced by the agency in order to produce the percentage figures. As a result, the population of Queensland is not necessarily considered an information asset for the Department of Child Safety. |
| 1. An information asset that is exchanged, received or commissioned externally should be identified as an information asset of both agencies | 1. An information asset exchanged between agencies on a regular cycle is to be treated as an agency information asset. The same is true for information that is commissioned by an agency and involves a third party external provider. 2. For example, Queensland Transport provides Queensland Police Service with driver licence information. Queensland Police Service should treat driver licence information as an information asset – as it is actively used in critical business processes of Queensland Police Service. 3. When Queensland Transport commissions a household travel survey, the results of the survey become an information asset of Queensland Transport, regardless of the entity that conducts the survey. |
| 1. Information present in two different content types results in two information assets | 1. Information may manifest as different content types and each of these types should be controlled as an asset. If not, then there is a risk that some assets will not be identified and insufficient controls applied to them. 2. It is important that information assets are identified by content type to allow assets to be assessed during ICT planning. This is particularly crucial as the business impact, future business value and technical condition of a transactional asset may differ from the analytical asset. 3. For example, Queensland Health maintains a patient register held as transactional content in a database. A summary of patient demographics is created from this source and stored in a data warehouse for reporting purposes. Queensland Health should consider these to be separate information assets: a transactional information asset containing individual patient records and details of patients; and an analytical information asset containing the summary information. 4. When naming similar assets across different content types, it is useful to apply a format noun which reflects the type of content. The following suggested format nouns best reflect the four content types*[[1]](#footnote-2)*: 5. **Transactional**: register, list, catalogue, inventory, dataset, index. 6. **Analytical**: dynamic report, result, hierarchy, tabulation, matrix. 7. **Authored**: feedback, content map, imagery, correspondence, message. 8. **Published**: static report, catalogue, content, library. |
| 1. An information asset should represent a collection of information | 1. An information asset is not a single data or information item. It should represent an identifiable collection of data, where a collection is a set of like or related information recognisable to the business users. 2. The use of terms such as data, record or detail in an asset name may indicate that the information asset is not a collection of like information but an instance within the collection. 3. For example, naming an information asset a patient register implies a collection containing patient names, addresses and other contact details. This is acceptable. Naming an information asset patient record could imply a single instance of one patient’s record details. This would not generally be an acceptable name for an information asset. |
| 1. An information asset should be recorded if the status of the information set remains unclear | 1. Even with the best guidelines, areas of ambiguity and doubt about the status of agency information will exist. 2. Ultimately, if an agency is in doubt about the status of a particular set of information, it is recommended to treat the set of information as an asset and include it in the information asset register. 3. Subsequent planning rounds will then allow for further refinement of the reported details. |

# Classification of an information asset

After an information asset is identified and named using the guidelines above, it should then be classified according to the Information classification framework. This involves classifying each information asset to a primary information domain, with the option to further classify the information asset to up to three secondary information domains.

## Using the Information classification framework

The Information classification framework is comprised of concepts known as domains. The domains in the framework are arranged in a hierarchy of preferred classifications, based on the following three levels:

**Level 1** – the broadest classes of information assets recognised by the QGEA. The level one domains describe 12 subject areas as shown in Figure 1 on page 9.

1. 

Figure 1 – Information Classification Framework level 1 domains

**Level 2** – major types of information assets within each broad domain.

**Level 3** – describes the more specific information asset domains (subjects or topics).

**Level 4** – variants that provide a pointer for the correct level 3 classification



**Level 1 domain**

**Level 2 domain**

**Level 3 domain**

An extract from the Information classification framework is provided in figure 2 below as an example.

Figure 2 – Extract from the Information Classification Framework

## Classifying information assets

The following guidelines are recommended to ensure a consistent approach to the application of the Information classification framework. This in turn will help to maintain the integrity of any analysis of an information asset classified to a particular domain.

| 1. Guideline | 1. Discussion |
| --- | --- |
| 1. Use either the level 4 domains or the variant list to classify the information asset name to the correct level 3 domains | 1. In addition to the domain levels (1, 2 and 3), the Information classification framework contains a number of level 4 domains and a variant list which is included to provide pointers from agency terminology to the correct Information classification framework domain. This is a similar concept to a library catalogue which includes ‘see’ and ‘see also’ terms. 2. The first step in classifying an information asset is to search for the name of the asset within the Information classification framework domains and variants. 3. For example, if an agency has a client register information asset, a search for ‘client’ within the Information classification framework reveals the domain client to be a type of commerce role. 4. In addition, the client register also contains information about people and companies. A search for ‘person’ reveals the domain person. A search for ‘company’ reveals public and proprietary companies which are a variant of a level 3 domain called organisation. Because a ‘client’ can be both a person and an organisation, it may be appropriate to classify the client register to a level 2 domain i.e. party. 5. Also, the client register may contain additional information about contact details, including mail but not email addresses. A search reveals two level 3 domains relating to addresses – electronic and geographic. Because the client register stores mail addresses only, the correct classification is the level 3 domain address. The fact that the addresses are only for mail-based contact allows us to also identify the address usage as postal. If the addresses had multiple usages then we may choose simply address usage. 6. The client register is primarily an information asset that is used by the business to identify clients and the means to contact these clients. It may therefore be appropriate to classify the Client Register information asset in the following manner:  * party (primary classification) * address/geographic (secondary classification) * role/commerce role (secondary classification) * address usage/postal (secondary classification). |
| 1. Identify the domains from the main business activity related to the information asset | 1. In cases where a search using the information asset name does not reveal a clear link to a domain, the business processes which act on the information asset may provide direction. Often, secondary domains can also be identified through analysis of other business processes or activities that use the data once it is captured. 2. For example, business users may refer to a register of carers which is used by a business process currently called register client details. A search using the word ‘carer’ may reveal no matches, whereas a search for ‘client’ using the business process name returns the commerce role domain. |
| 1. Identify domains through common associations and logical information groupings | 1. An information asset is identified by the logical grouping or collection of ‘like’ information. Identification of domains will also benefit from the same analysis of the common or logical collections often present in business information. 2. For example, a facility/building information asset probably also contains address details. Therefore, identification of a primary domain for building should lead to an analysis of the asset for common associations of this type such as address. |
| 1. An information asset should have a dominant primary information domain | 1. An information asset, if defined at an appropriate level of granularity, should have a single dominant concept or information kernel upon which it is based. Often this will be the subject noun in the name of the information asset. In cases where the identification of a primary domain for the information asset is difficult, it may indicate that the asset is too large and needs to be decomposed into smaller, more manageable information assets. |
| 1. An information asset should not have excessive secondary information domains | 1. An information asset, if defined at an appropriate level of granularity, should not have many more secondary domains beyond the current limit of three per asset. If a large number of secondary domains are found, it may indicate that the asset is too large and needs to be decomposed into smaller, more manageable information assets. 2. **Note:** The use of three secondary domains per information asset is the recommended limit and the maximum for baseline reporting purposes. In some cases, agencies may wish to classify information assets to more than three secondary domains, so the definition of ‘excessive’ secondary domains is an agency decision. |
| 1. An information asset rarely has only one domain classification | 1. When an information asset is supported by only one information domain, the asset may be too narrowly scoped. This is particularly relevant as some information types on their own may offer little capability to support business processes, assist in making decisions or add business value. 2. For example, it would be rare to have an address list as a separate asset (exceptions may include Department of Environment and Resource Management). 3. **Note:** Many agencies will have address details as a secondary aspect to a customer list or are more interested in geo-spatial locations rather than strictly geographic addresses. |
| 1. An information asset often contains like information archetypes | 1. The Information classification framework provides a broad categorisation of its component information domains according to three information archetypes: motivators, moments and entities[[2]](#footnote-3). An information asset may be too large if its classifications cross over the information archetype boundaries. Although it is acceptable for the information asset to be classified in more than one archetype, an agency may find that it is more common for an asset to contain multiple domains from within the same archetype. |
| 1. Information domains also provide a mechanism for naming an information asset | 1. The Information classification framework can assist in not only verifying the correct level of granularity for information assets, but also in providing assistance during the naming of assets. 2. For example: ticketing equipment inventory provides clarity on the domain (i.e. equipment). 3. **Note:** This is particularly useful when identifying information assets from within applications. One approach is to perform an initial mapping of an application to the domains. From this an agency can then break down the application into logical information collections using the domain mappings. |
| 1. An information asset is only classified using products or services when it describes a product or service, or represents an instance of a product or service | 1. An information asset should only map to the products or services domains when it contains information describing the product or service, or contains an instance of products or services received. That is, the information asset represents a catalogue, overview, inventory or definition of products or services. 2. For example, the Department of Public Works (DPW) home page ([www.publicworks.qld.gov.au](http://www.publicworks.qld.gov.au/)) contains content which defines the major services of the department. The content could be mapped to the service domains applicable to DPW as it contains information describing the services of government. In Queensland Transport, a vehicle register may contain the records of individual vehicle registrations issued by the department. This may be mapped to the registration domain. |
| 1. An information asset is classified based on the content of information assets, not its form or relationships | 1. The purpose of the Information classification framework is to classify information assets based on the information they contain. That is, the subject or information type. However, in some cases the content, the purpose for collection or even the form (structure) may appear similar. It is important to identify the true nature of the information held within an information asset. 2. A good question to ask is, ‘If I could read out loud the information contained in this asset, what would an independent listener say it was about?’ 3. For example, many information assets take the form of documents (electronic or physical). However, the information in these documents is not about documents themselves but about other subject (domains). 4. For example, a student enrolment register may logically be associated with a contract for education services, but it may not contain any contractual details. Therefore, it would be incorrect to classify this asset using the responsibilities/contract/service domain; it should be classified as person. |

# Application specific considerations

1. There are two common agency ICT applications that need specific considerations when identifying information assets, namely, eDRMS and web content management systems.

## Electronic document and records management systems

1. When dealing with an electronic document and records management system (eDRMS), it is tempting to identify information assets that relate to the content of the documents and records held within the system. This is especially true when the eDRMS holds complete electronic copies of the source documents.
2. However, for the purpose of ICT planning, an eDRMS is considered to contain information about the tracking of files and documents. The contents of any documents or files it holds are to be dealt with separately. If this limitation is not applied, an eDRMS could potentially map to the majority of domains within the classification framework. Such a scenario would provide little value for future analysis.
3. A more valuable approach is to identify an eDRMS as containing information assets such as business records repository or file and records inventory (or similar names according to each agency’s terminology). It is then recommended that agencies classify these against the resources/artefact domains as primary, with the option to record secondary domains such as interactions/action/correspond and interactions/action/transact.
4. The situation with an eDRMS may also manifest in other similar pure catalogue, tracking or inventory systems, such as library cataloguing. If in doubt, agencies are encouraged to seek assistance from the Queensland Government Chief Information Office.

## Static website content and web content management systems

1. As with eDRMS, agency websites and associated web content management systems (WCMS) pose their own unique challenges for information asset identification and subsequent classification.
2. For the purpose of ICT planning, it is recommended that static website content is decomposed into information assets that match the site (navigation) structure (or division of information within the web content management system). Each of these main areas, once treated as an asset, can then be classified. When classifying information assets based on web content, it is suggested that agencies use the product or service domains as the primary domain. Secondary domains will typically include things such as controls/ administrative/policy and plans/direction/strategy.
3. If a website contains a dynamic application or is the gateway for an application (such as VCheck which is accessed over the web via [www.transport.qld.gov.au](http://www.transport.qld.gov.au)) then the dynamic component or application may generate its own set of information assets.

# Information asset identification and classification form

The [Information asset identification and classification form](https://www.qgcio.qld.gov.au/products/qgea-documents/570-workflow/3754-identification-and-classification-of-information-assets-form) provides a single checklist for identifying information assets. Agencies are encouraged to apply this checklist when executing the ICT planning methodology. It can be used during analysis or as a quality assurance mechanism before finalising ICT baseline submissions. Agencies may supplement this checklist with their own specific criteria, if required.

1. Note that format nouns are not mutually exclusive to a particular content type, however some format nouns have a closer association to certain content types than others. These suggestions are not exhaustive and agencies are encouraged to develop their own format nouns as appropriate. [↑](#footnote-ref-2)
2. The archetypes are defined in the *Glossary for ICT Planning Methodology.* [↑](#footnote-ref-3)