NATIONAL ARCHIVES OF AUSTRALIA

Developing a Functions Thesaurus

Guidelines for Commonwealth Agencies

July 2003

ISBN 1 920807 06 3

CONTENTS

INTRODUCTION **6**

Acknowledgments 7

Overview of steps to create a functions thesaurus 7

Figure 1 – Overview of functions thesaurus development process 8

1 DETERMINING RELATIONSHIPS BETWEEN TERMS **9**

1.1 Relationships in a thesaurus 9

Table 1 – Relationships established in a thesaurus 9

Table 2 – Definitions of relationships established in a thesaurus 9

1.1.1 Hierarchical relationship: broader terms (BT) and narrower terms (NT) 10

Figure 2 – Relationships between broader and narrower terms 11

1.1.2 Association relationship: related terms (RT) 11

Table 3 – Example of related terms 12

1.1.3 Equivalence relationship: 'USE' and 'Use For' terms 12

Table 4 – Descriptions of types of non-preferred terms 12

Table 5 – Examples of preferred and non-preferred terms 12

Table 6 – Examples of preferred and non-preferred terms 14

1.2 Sources for non-preferred terms 14

Exercise 1 15

2 CONSTRUCTING A FUNCTIONS THESAURUS **16**

2.1 Constructing a thesaurus of core functions 16

2.1.1 Use of generic activities from Keyword AAA (CV) 16

2.1.2 Applying a source to terms 17

2.1.3 Steps in constructing a thesaurus of core functions 17

Table 7 – Step-by-step process for constructing a thesaurus of core functions 17

2.2 Constructing a merged functions thesaurus 19

2.2.1 Using Keyword AAA (CV) in a merged functions thesaurus 19

2.2.2 Considerations when merging with Keyword AAA (CV) 19

2.2.3 Steps in constructing a merged functions thesaurus 19

Table 8 – Step-by-step process for constructing a merged thesaurus 20

Figure 3 – Main steps in creating a merged functions thesaurus 22

2.3 Common errors in the relationships established between terms 23

Table 9 – Identifying problems and solutions with thesaurus relationships 23

Exercise 2 24

3 MANAGING CHANGE TO A FUNCTIONS THESAURUS 25

3.1 Reasons to modify a functions thesaurus 25

3.1.1 Managing modifications to a functions thesaurus 25

3.1.2 Documenting changes to the thesaurus 26

3.2 Modifying a functions thesaurus when gaining a function 26

3.2.1 The business classification scheme 26

3.2.2 Incorporating terms for a new function into the thesaurus 26

Figure 4 – Tasks to integrate an inherited function into a merged thesaurus 28

3.3 Modifying a functions thesaurus when losing a function 29

3.3.1 Modifying a thesaurus when a function is sold to private enterprise 29

3.4 Keeping terminology current in a functions thesaurus 30

3.4.1 Reasons to update terms and relationships 30

3.4.2 Requests for changes from users 30

3.4.3 Amending terms in the functions thesaurus 30

Figure 5 – Process to amend terms in the functions thesaurus 32

3.5 Considerations when a function is outsourced 33

3.5.1 Managing classification and titling with a service provider 33

Exercise 3 33

APPENDIXES 34

A Thesaurus change suggestion template 34

Figure 6 – Suggested template for a thesaurus change form 34

B Selection criteria for personnel to develop and maintain a thesaurus 35

Recordkeeping competencies 35

Management competencies 36

Thesaurus content construction skills 36

Thesaurus software application competencies 37

C Notes on thesaurus software applications 38

Functional requirements for thesaurus software 38

Business rules for thesaurus validation and data entry 39

Options for users to select terms for classifying records 39

Using a thesaurus browser application 40

Customising the display of terms 40

Retrieval features 40

Using a thesaurus in the query and retrieval process 41

D Subject and other thesauruses 42

Criteria for using a subject thesaurus for classification and retrieval 42

Integration of subject thesauruses into retrieval tools 42

E Answers to the Exercises 44

F Glossary 46

G References 49

INTRODUCTION

These guidelines give practical advice to Commonwealth agencies wanting to develop a functions thesaurus to support records management processes in their organisation. The management of records and other business information can be greatly facilitated by developing and implementing a functions thesaurus to help classify, title, retrieve, sentence and dispose of records.

These guidelines complement the *Overview of Classification Tools for Records Management*, which provides advice on the types of classification tools available for records management and describes the characteristics of each. The benefits of a functions approach, the business environment in which a tool is to work, and the stakeholders in the development and use of a tool are also examined. The *Overview of Classification Tools* is essential reading for making decisions on which classification tool is the most appropriate for an organisation and for undertaking any project to develop a classification tool.

Because a functions thesaurus is the most sophisticated classification tool examined in the *Overview of Classification Tools*, and the most resource intensive to develop, we have developed these additional guidelines. Nonetheless, some of the principles relating to the development and maintenance of a functions thesaurus can also apply to other classification tools.

These guidelines apply once you have decided that a functions thesaurus is the type of classification tool required by your organisation. They assume that your organisation has developed a business classification scheme in accordance with the methodology outlined in 'Step A – Preliminary investigation' and 'Step B – Analysis of business activity' of DIRKS: A Strategic Approach to Managing Business Information.[1] In many cases there is a distinct advantage in completing 'Step C – Identification of recordkeeping requirements' of the DIRKS methodology before commencing work to develop a functions thesaurus. Step C will often expose recordkeeping requirements that have been inadequately addressed by the organisation. Consequently, a functions thesaurus will usually require revision if developed before completing Step C.

Divided into three parts, the guidelines progress through the development of a functions thesaurus, describing step-by-step the process of constructing both a thesaurus of core agency functions and a merged thesaurus, incorporating administrative functions with core agency functions to form an integrated classification tool. Each part includes a practical exercise, with answers, to help you put concepts into practice.

The guidelines include references to the Australian Standard for Records Management, AS ISO 15489 – 2002,[2] which supersedes AS 4390, and to *Keyword AAA: A Thesaurus of General Terms, Commonwealth Version* (May 2001).[3]

The development and management of a functions thesaurus requires specialised training and skills. Appendix B of these guidelines advises on desirable competencies, skills and knowledge for staff developing and maintaining a functions thesaurus.

---

[1] See www.naa.gov.au/recordkeeping/dirks/dirksman/dirks.html

[2] See www.naa.gov.au/recordkeeping/rkpubs/advices/advice58.html

[3] See www.naa.gov.au/recordkeeping/control/keyaaa/summary.html – hereafter this document is referred to as Keyword AAA (CV)

These guidelines are a revision of an earlier publication, *Developing a Functions Thesaurus: Guidelines for Commonwealth Agencies*, issued in 2000.

The current guidelines have been issued as an interim publication, pending the outcomes of work of the Standards Australia IT-021-09 Committee, convened to develop a Records Management Classification best practice standard. The National Archives of Australia is actively involved in the committee's work on applying classification principles to a functions-based approach and translating these into implementation tools. These guidelines will be reviewed in the light of the committee's work.

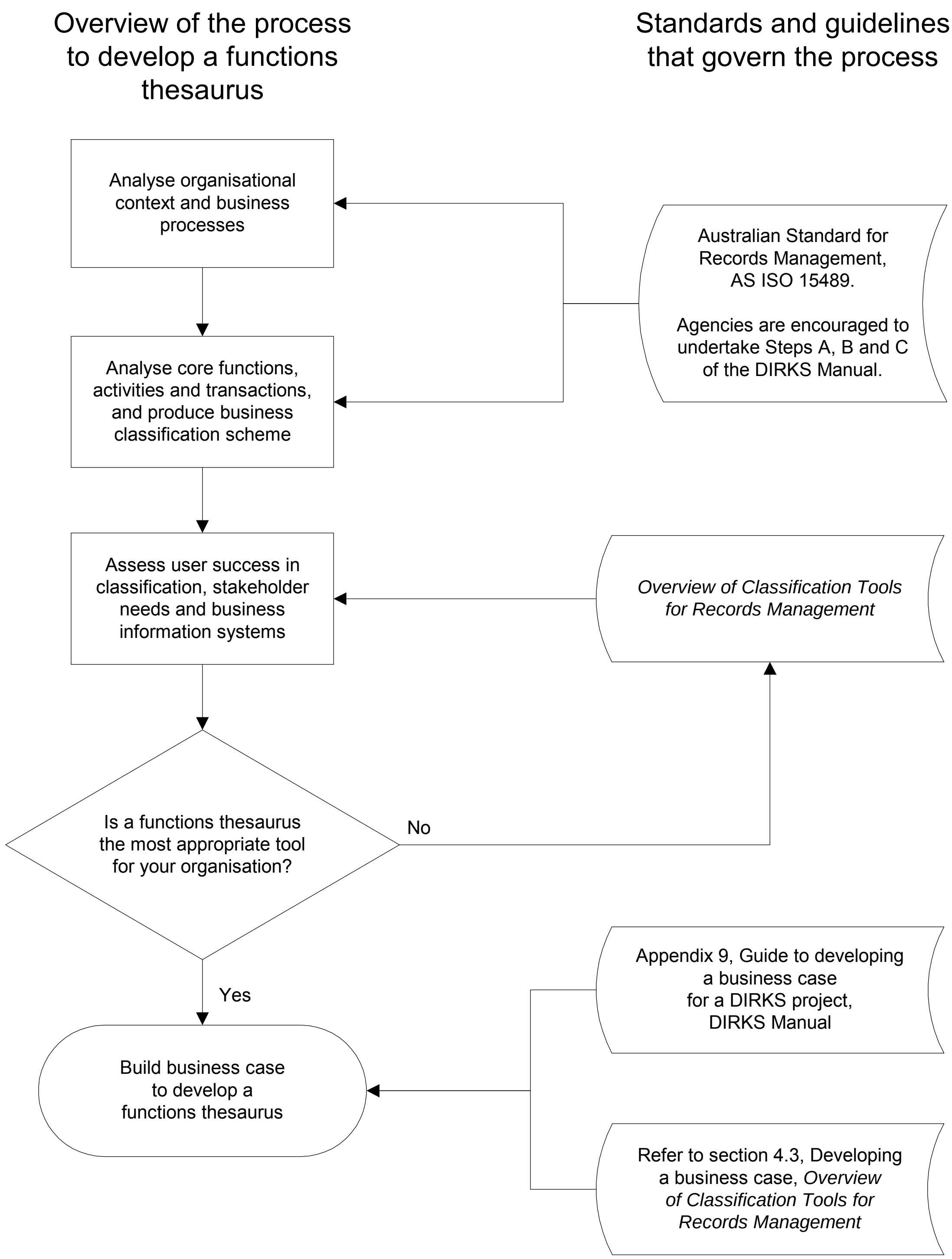
Acknowledgments

The National Archives of Australia would like to express its appreciation to the Attorney-General's Department for permission to use examples derived from its business classification scheme and to the State Records Authority of New South Wales, which has assisted in the development of this material.

Overview of steps to create a functions thesaurus

Figure 1 represents the process of developing a functions thesaurus. It includes essential prerequisite steps and refers to standards and guidelines used during the process.

Figure 1 – Overview of functions thesaurus development process



1 DETERMINING RELATIONSHIPS BETWEEN TERMS

A functions thesaurus is a classification tool to assist organisations to manage their records and other business information. It is:

> a controlled list of terms linked together by semantic, hierarchical, associative or equivalence relationships. Such a tool acts as a guide to allocating classification terms to individual records.[4]

A thesaurus is designed to enable users to identify preferred, or authorised, terms for classifying and titling records and to offer a range of paths to reach these terms.[5] It also facilitates strategies for retrieving records and reduces the likelihood of an unsuccessful search, or one resulting in a misleading or irrelevant outcome. This functionality is achieved by establishing paths between terms. This section explains the paths, or relationships, that are established between terms in a thesaurus.

1.1 Relationships in a thesaurus

A thesaurus is based on three main relationships.

Table 1 – Relationships established in a thesaurus

Hierarchy	Establishes a relationship between broader and narrower concepts
Association	Establishes a relationship between related concepts at the same level
Equivalence	Establishes a relationship between preferred terms – ie, terms that are to be used (also called authorised terms) – and non-preferred terms – ie, terms that are not to be used (also called unauthorised terms).

Each relationship between terms in a thesaurus is displayed through particular conventions, expressed by an abbreviation. These abbreviations and their definitions, include the following.

Table 2 – Definitions of relationships established in a thesaurus

Type	Name	Abbreviation	Definition
Hierarchy	Broader term	BT	• Indicates a concept that has a wider meaning relative to another term. • Terms that are superior in a category.
Hierarchy	Narrower term	NT	• Indicates a term with a more specific meaning relative to a broader category. • Subordinate component of a wider category.

[4] Australian Standard for Records Management, AS ISO 15489, Part 2, Clause 4.2.3.2.

[5] AS ISO 15489, Part 2, Clause 4.2.3.2.

Type	Name	Abbreviation	Definition
Association	Related term	RT	• Indicates similarities with other concepts. • Indicates a term is associated with another term but is not a synonym, nor is it in a hierarchical relationship.
Equivalence	Use for	Use For or UF	• Indicates a term has a preferred status. • Indicates non-preferred terms associated with the preferred term – such as synonyms, or abbreviations – which are not to be used.
Equivalence	Use	USE	• Indicates a term has a non-preferred status. • Indicates the preferred term to use when a non-preferred term has been selected.

1.1.1 Hierarchical relationship: broader terms (BT) and narrower terms (NT)

This relationship is based on a ranking from a superior to a subordinate position. Terms that are superior in a category are at a superordinate level and terms that fall under or below a category are at a subordinate level. The broader term represents a class or a whole, while the narrower term refers to components or parts of the broader concept.[6]

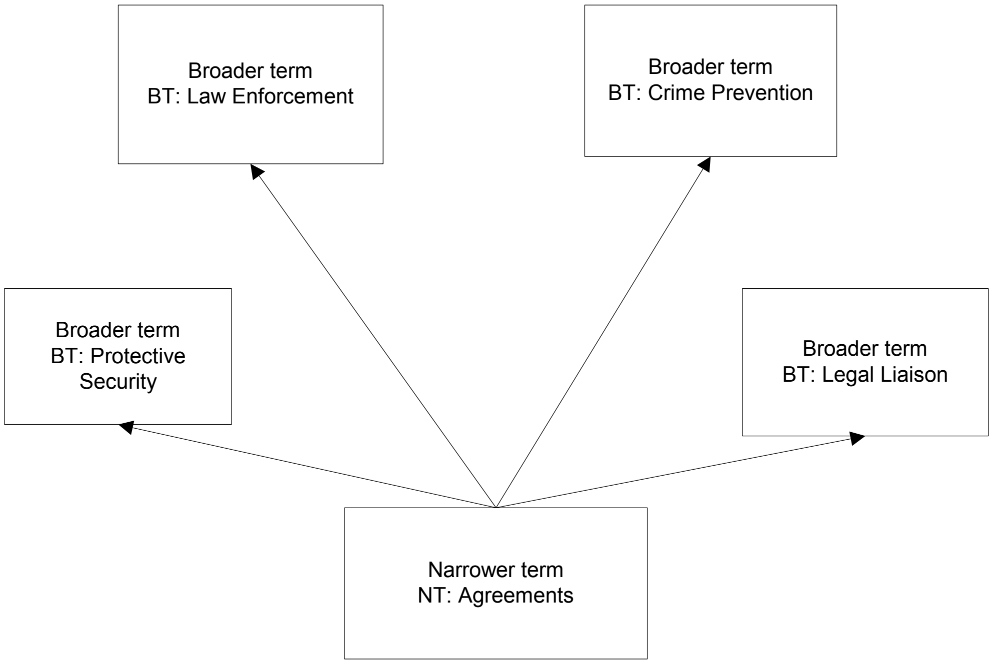
The relationship between the terms is expressed by the following abbreviations:

- BT (broader term) – written as a prefix to the superordinate term
- NT (narrower term) – written as a prefix to the subordinate term

A functions thesaurus extends the one-to-one relationships established in a business classification scheme to one-to-many relationships, within a hierarchy of terms. Figure 2 demonstrates the one-to-many relationships of the activity of 'Agreements' to a number of functions undertaken by the Attorney-General's Department.

[6] ISO 2788 – 1986, Clause 8.3.

Figure 2 – Relationships between broader and narrower terms



Source: Attorney-General's Department, 2003

1.1.2 Association relationship: related terms (RT)

An association relationship is established to indicate that a term has similarities with other concepts. A related term relationship alerts users to the fact that other records and information of interest may be classified under a different, but related, set of terms. Specifying a related term relationship is useful when the definition for a term does not seem to quite fit. Related term relationships are only established between preferred terms.[7]

A related term relationship is not intended to accommodate variations in expression. Related terms usually refer to a single term in the thesaurus. If there are associations between combinations of terms, these are best explained in a scope note. Further information on developing scope notes is available in Step B of the DIRKS Manual.

The association relationship is not hierarchical. It is a thesaurus convention that terms with this kind of relationship are at the same hierarchical level. That is, activities are related to other activities; activities *cannot* have an RT relationship to functions.

In the example in Table 3, when searching using the function-level term 'Legal Counsel', records of the function of 'Legal Services' are also retrieved.

[7] ISO 2788 – 1986, Clause 8.4.

Table 3 – Example of related terms

Level in thesaurus hierarchy	Term	Related term
Function	Legal Counsel	**RT**: Legal Services
Activity	Child Maintenance Management	**RT**: Child Custody and Access Orders Management

Source: Attorney-General's Department, 2003

1.1.3 Equivalence relationship: 'USE' and 'Use For' Terms

Equivalence relationships are established between synonyms, which in many contexts have meanings that are considered virtually interchangeable. However, in a controlled vocabulary, such as a thesaurus, where meanings are deliberately restricted, only one term is considered the preferred term. Synonyms act as pathways to preferred terms.[8]

'USE' and 'Use For' prefixes identify the relationship between preferred and non-preferred terms, when two or more terms may be regarded as referring to the same concept.

The nature of the relationship is expressed by the following conventions:

- **USE** – which is written as a prefix to the preferred term
- **UF** or **Use For** – which is written as a prefix to the non-preferred term.

The power of a thesaurus as a retrieval tool stems from establishing relationships between preferred and non-preferred terms. For example, a record titled 'Fleet Management – Acquisition – Utility' will not be retrieved with a search using term 'Ute', unless 'Ute' is established as a non-preferred term for 'Utility':

Ute – **USE**: Utility.

By establishing this relationship the desired record will be retrieved.

There are two main types of non-preferred terms that are added to a thesaurus to assist navigation to the preferred term but which are not and must not be used to classify a record: synonyms and variations in description.

Table 4 – Descriptions of types of non-preferred terms

Synonyms	Terms that are considered similar in meaning, or similar ways of expressing the same concept. These can also include variations in spelling, combinations of terms and singular and plural versions. One of these terms is selected as the preferred term and the remaining terms become non-preferred terms. Non-preferred terms will navigate users to the preferred term.
Variations in description	Terms that also describe functions, activities and topics, perhaps reflecting previous agency classification and titling conventions, but which are non-preferred terms.

[8] ISO 2788 – 1986, Clause 8.2.

Table 5 gives examples of preferred and non-preferred terms. The distinction between the two types of non-preferred terms is presented for illustration only. Such distinctions may be made explicit during the development of the thesaurus but are not explicit in the published tool.

Table 5 – Examples of preferred and non-preferred terms

Preferred term USE	Non-preferred term Variations in description	Non-preferred term Synonym
USE: Child maintenance management	**Use For**: Maintenance orders	**Use For**: Maintenance
USE: Appointment termination	**Use For**: Cessation of appointment	**Use For**: Removal from office
USE: Advice	**Use For**: Submissions	**Use For**: Opinions
USE: Legal instruments	**Use For**: Instruments of authority	**Use For**: Ordinances

Source: Attorney-General's Department, 2003

In theory, there is no limit on the number of non-preferred terms in a functions thesaurus. Limitations are imposed by factors such as resources. If users are used to free-text searching, they will expect to be able to enter terms with which they are familiar. Obsolete terms, used in the past, should be included as non-preferred terms.

For example, if you have a current term of 'Information Management', include previous terms, such as 'Registry', 'Records Management' and 'Library' in the thesaurus as non-preferred terms. Familiar terms operate as signposts, leading users to the preferred term. In this way, users are encouraged to accept and adopt the thesaurus:

Registry	**USE**: Information Management
Records Management	**USE**: Information Management
Library	**USE**: Information Management
Information Management	**Use For**: Registry
	Use For: Records Management
	Use For: Library

Table 6 shows how preferred terms from Keyword AAA (CV) can be identified by using non-preferred terms. It does not matter that these extra terms have a 'subject' flavour, providing they guide users to a preferred term.

Table 6 – Examples of preferred and non-preferred terms

Keyword AAA (CV) Preferred Term	Keyword AAA (CV) Non-preferred Term	Agency Non-preferred Terms
Marketing	**Use For**: Public image	**Use For**: Branding **Use For**: Promotion
Fleet Management	**Use For**: Transport	**Use For**: Car pool

1.2 Sources for non-preferred terms

There are many avenues for gathering non-preferred terms. However, you will need to consider the resources you have available to gather the terms and determine their relevance. The main sources include:

- search engine logs – both intranet and Internet logs will reveal users' search patterns and words that are spelt incorrectly. These will also indicate failed searches, searches that returned a large set of results (possibly without enough specificity for the searcher) and searches with a low number of results (which may indicate a more precise return of results)
- organisation-wide publications such as strategic plans and annual reports. Some of these would have been examined to identify business functions during the development of the agency's business classification scheme. A list of non-preferred terms may have emerged in the process. Indexes to annual reports can also have useful terms to enter into the thesaurus
- examples of previous record titling practices
- discussions with users on naming conventions – word association exercises can produce valuable terms for the thesaurus.[9] In such exercises you could start with a term from the business classification scheme and ask for other words that spring to users' minds.

An example of a full business classification scheme and related functions thesaurus is provided in Keyword AAA (CV), which most agencies have access to through licencing arrangements with the National Archives.

Many more relationships are displayed in a functions thesaurus than in a business classification scheme, demonstrating how the functionality and accessibility of information is enhanced in a thesaurus structure.

[9] Maria Lopez-Huertas, 'Thesaurus structure design: A conceptual approach for improved interaction', *Journal of Documentation*, 1997, p. 171.

Exercise 1

Examine the following pairs of relationships. Are they true or false? If they are false, what would be a true relationship? The examples given in this exercise are sourced from Keyword AAA (CV).

Example	Classification terms	Thesaurus terms	True	False
1	Legal Services Litigation	**NT**: Litigation **BT**: Legal Services		
2	Court Cases Litigation	**USE**: Litigation **USE**: Court Cases		
3	Litigation Claims	**RT**: Claims **NT**: Litigation		
4	Legal Disputes Litigation	**USE**: Litigation **Use For**: Legal Disputes		

2 CONSTRUCTING A FUNCTIONS THESAURUS

A functions thesaurus can encompass only the core functions the organisation performs or it can describe the entirety of functions performed by the organisation. Core functions are responsibilities unique to the organisation. A number of other responsibilities are performed by an organisation that are common to most other Commonwealth agencies. These functions are referred to as administrative functions. The set of terms for core and administrative functions can be maintained as separate tools or can be combined into one product, called a merged thesaurus. Together they provide the framework for classification and control of all the organisation's records and business information.

2.1 Constructing a thesaurus of core functions

The primary purpose of a functions thesaurus is to navigate users to terms that can be used for classifying, titling and retrieving records and other business information. Constructing a functions thesaurus involves creating logical relationships between business functions, activities and topics. In thesaurus language, it involves relating terms through broad, narrow, related and equivalence relationships, and linking non-preferred terms.

At this point, the number of classification levels to incorporate into the thesaurus will have been decided, topic level terms developed and the rules – or conventions – for titling established. *Overview of Classification Tools* gives detailed advice on the development of levels of classification and aligning the classification tool to the business classification scheme and to the disposal authorities relevant to the organisation.

The majority of work in constructing a functions thesaurus is in linking the equivalence relationships – that is, the relationships between preferred and non-preferred terms. Some of the broader and narrower relationships have already been established between functions and activities, through the hierarchical structure of the business classification scheme. The association relationships between terms, relating similar concepts at the same level, also need to be added.

Scope notes will have been developed for the function and activity terms in the business classification scheme. Some of these scope notes will be original, developed explicitly by the organisation, to describe its core functions. Further information on developing scope notes is available in Step B of the DIRKS Manual. Other terms and their descriptions can be found in generic functions thesaurus tools, such as Keyword AAA (CV).

2.1.1 Use of generic activities from Keyword AAA (CV)

Keyword AAA (CV) is a functions thesaurus for records classification developed to describe administrative functions performed in most Commonwealth agencies. It does not cover core agency functions. Keyword AAA (CV) has been developed to three levels of classification. Relationships between terms at all levels are defined in the product.

Even though Keyword AAA (CV) does not describe core functions, it is likely that some Keyword AAA (CV) activity descriptors have been included in the organisation's own business classification scheme. It is useful to look through the thesaurus section of Keyword AAA (CV) and note the range of relationships established with these terms and whether these are appropriate for use in the organisation's customised thesaurus. The use of third-level classification in Keyword AAA (CV) is also worth evaluating.

The rules for using Keyword AAA (CV) terms in an agency's business classification scheme are explained in the DIRKS Manual at 'Appendix 6 – Practical advice for using Keyword AAA and AGIFT terms'.

2.1.2 Applying a source to terms

When using a thesaurus software application, terms usually need to be labelled with their source during the process of keying in terms. Keyword AAA (CV) terms should be clearly tagged in your functions thesaurus with Keyword AAA (CV) as the source. On occasions, terms may be sourced from the Australian Governments' Interactive Functions Thesaurus (AGIFT)[10] where there is an appropriate term. Unless an agency undertakes a whole of government function, it is unlikely a term from AGIFT will be specific enough for record management purposes, but terms will correlate from time to time. These terms should be tagged with 'AGIFT' when keyed into the functions thesaurus software. Core functions and activities developed by the organisation should be tagged with something like 'Agency' or 'Organisation' as the source.

2.1.3 Steps in constructing a thesaurus of core functions

There are general steps that apply when creating a functions thesaurus. Following the steps below will ensure consistency and avoid conflicts in relationships between terms.

Constructing a functions thesaurus is much easier if you use a functions thesaurus software application. Appendix C: Notes on Thesaurus Software Applications advises on aspects of thesaurus application functionality and desirable features.

Table 7 – Step-by-step process for constructing a thesaurus of core functions

Step	Actions
1	• Enter core function terms from the business classification scheme. Ensure these are tagged as preferred terms. • Enter any additional function terms devised to accommodate user requirements. Tag these as preferred terms. • Insert scope notes next to function term. • Acknowledge the source of the term, which will be your organisation-specific classification scheme. • Ensure terms sort into alphabetical order.

[10] See www.naa.gov.au/recordkeeping/gov_online/agift/summary.html

Step	Actions
2	• Enter core activity terms from the organisation's business classification scheme, and ensure these are tagged as preferred terms. • Enter additional activity-level terms developed to accommodate user requirements (eg 'case' and 'project' types of 'mega-activities'), and ensure these are tagged as preferred terms. • Insert scope notes next to activity term. • Acknowledge the source of the term, for example Keyword AAA (CV) or your organisation-specific business classification scheme. • Ensure terms sort into alphabetical order.
3	• Insert the terms formulated for the third level of classification. These topic terms may relate to groups of transactions, may include subject concepts, record types or abbreviations used in the organisation. Ideally these terms should not be more than two words long. • Ensure these terms are tagged as preferred terms. • If a scope note is required to explain the term, insert it adjacent to the term. • Acknowledge the source of the term. • Ensure terms sort into alphabetical order.
4	• Establish broader and narrower relationships. These relationships will occur simultaneously in most thesaurus applications. • Ensure function terms have a BT relationship with each activity term undertaken to perform the function. • Ensure activity terms have an NT relationship to the function terms they are attached to. • Ensure activity terms have a BT relationship to each topic term undertaken to perform the activity. • Ensure topic terms have an NT relationship to each activity term they relate to.
5	• Establish associative relationships with an RT tag for terms at the same level. For example, a function term can only have an RT relationship with another function term and an activity term can only have an RT relationship with another activity term. A topic term can be related to another topic term. • These RT relationships will occur simultaneously in most thesaurus applications.
6	• Enter non-preferred terms. These are variations to the preferred terms, outdated terminology, spelling variations etc that are used in the organisation to describe functions, activities and topics. These terms are put into the thesaurus but are not considered acceptable for use. • Indicate the equivalence relationship by linking the non-preferred term to the preferred term. Ensure preferred term has a 'USE' tag. • Ensure terms sort into alphabetical order.
7	• Ensure the non-preferred terms are entered at the preferred term. • Indicate the equivalence relationship by attaching a 'Use For' tag to the non-preferred term.
8	• Review the listing and ensure no illogical relationships have been created.

Constructing a thesaurus does not have to be conducted in strict accordance with the process outlined above. Constructing the thesaurus can begin at the topic level. The important point to remember is that relationships must be established according to thesaurus rules.

2.2 Constructing a merged functions thesaurus

A merged thesaurus is one that integrates the terms describing an organisation's core functions with the terms describing the common administrative functions also performed by the organisation. This creates a strong classification and retrieval tool, especially when applied in a business information system that is used extensively across the organisation to manage records and other business information relating to many of the organisation's functions.

2.2.1 Using Keyword AAA (CV) in a merged functions thesaurus

By itself, Keyword AAA (CV) is not an effective means of classifying all the records of an agency. It is designed to be used in conjunction with the set of terms that describe an agency's core functions and activities. The result is a merged thesaurus, which provides full coverage of the organisation's business activities. A merged thesaurus offers comprehensive, controlled vocabulary coverage, tailored to an organisation, in a single classification tool.

2.2.2 Considerations when merging with Keyword AAA (CV)

Some Keyword AAA (CV) activity terms may have already been incorporated into the organisation's business classification scheme and linked to its core functions. The effect of utilising all Keyword AAA (CV)'s terms, and how they will interact with the terms for the core functions and activities, will need to be determined at this stage.

In some cases, a term listed as an administrative function is undertaken by an agency as a core function. An example of this occurs in the Attorney-General's Department. The core functions of the organisation cover management of legal services in the Commonwealth. However, as an agency, it will also undertake the administrative function of 'Legal Services'. 'Legal Services' is a Keyword AAA (CV) term and, therefore, is required in a merged functions thesaurus. The Attorney-General's Department developed its own unique terms, and scope notes, to describe its core functions of managing legal services in the Commonwealth, during the development of its business classification scheme. The difference between its core terms, 'Legal Counsel', 'Legal Liaison' and 'Legal Case Management' and the Keyword AAA (CV) term of 'Legal Services' is clarified through scope notes.

2.2.3 Steps in constructing a merged functions thesaurus

The process to construct a merged thesaurus is explained in the steps below and then illustrated in the diagram that follows. It is recommended an electronic version of Keyword AAA (CV) is loaded into the thesaurus software first. Keyword AAA (CV) is available in a variety of formats that are suitable for transfer into other software packages.[11] Once this is done, agency-specific functions can be entered manually, or exported from other thesaurus tables in your business information system. Table 8 begins with the initial step of loading Keyword AAA (CV) and then continues with the

[11] See www.naa.gov.au/recordkeeping/control/keyaaa/summary.html

same eight steps outlined in the construction of a thesaurus of core functions, creating nine steps in the process to build a merged thesaurus. However, Step 7 in the process below has additional information, on the addition of non-preferred terms from Keyword AAA (CV).

If you are using a thesaurus software application, the steps in the process must be considered in light of its functionality. For example, in some applications all terms begin as preferred (or authorised) terms. Non-preferred terms are then tagged with an unauthorised status. In other applications, a separate action/button must be activated to indicate the term is authorised to classify and title records. Most thesaurus applications have the basic principle that links cannot be created until terms have been entered and saved. Test that preferred terms can be allocated to file titles and non-preferred terms cannot.

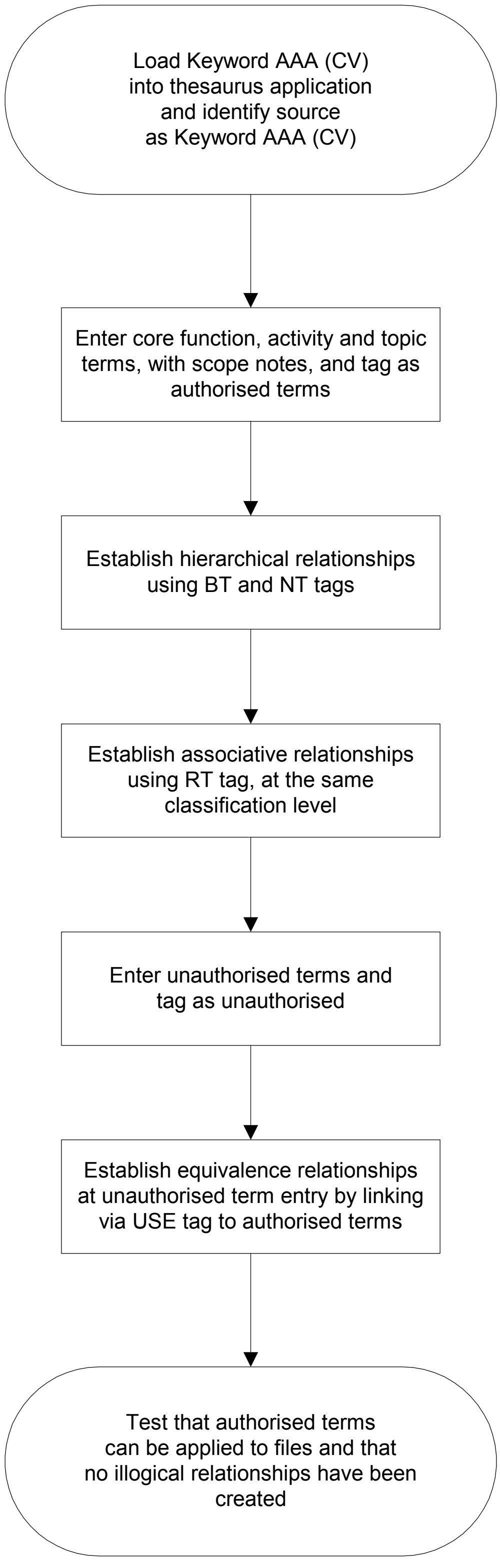
Table 8 – Step-by-step process for constructing a merged thesaurus

Step	Actions
1	• Load Keyword AAA (CV) terms into the thesaurus software. • Ensure preferred terms are tagged as authorised and non-preferred terms are tagged as unauthorised. • Note Keyword AAA (CV) as the source. • Ensure terms sort into alphabetical order.
2	• Enter core function terms from the business classification scheme and tag these as preferred terms. • Enter any additional function terms devised to accommodate user requirements. Tag these as preferred terms. • Insert scope notes next to function terms. • Acknowledge the source of the term, for example your organisation-specific business classification scheme or Keyword AAA (CV). • Ensure terms sort into alphabetical order.
3	• Enter core activity terms from the organisation's business classification scheme and ensure these are tagged as preferred terms. • Enter additional activity level terms developed to accommodate user requirements (eg 'case' and 'project' types of 'mega-activities') and ensure these are tagged as preferred terms. • Insert scope notes next to activity terms. Acknowledge the source of the term, for example Keyword AAA (CV) or your organisation-specific business classification scheme. • Ensure terms sort into alphabetical order.
4	• Insert the terms formulated for the third level of classification. These topic terms may relate to groups of transactions, may include subject concepts, record types or abbreviations used in the organisation. Ideally these terms should not be more than two words long. • Ensure these are tagged as preferred terms.

Step	Actions
	• If a scope note is required to explain the term, insert it adjacent to the term. • Acknowledge the source of the term. • Ensure terms sort into alphabetical order.
5	• Establish broader and narrower relationships. These relationships will occur simultaneously in most thesaurus applications. • Ensure function terms have a BT relationship with each activity term undertaken to perform the function. • Ensure activity terms have an NT relationship to the function terms they are attached to. • Ensure activity terms have a BT relationship to each topic term undertaken to perform the activity. • Ensure topic terms have an NT relationship to each activity term they relate to.
6	• Establish associative relationships with an RT tag. These terms will be at the same level of classification. For example, a function term can only have an RT relationship with another function term and an activity term can only have an RT relationship with another activity term. A topic term can be related to another topic term. • These RT relationships will occur simultaneously in most thesaurus applications.
7	• Enter non-preferred terms. These are variations to the preferred terms, outdated terminology, spelling variations etc that are used in the organisation to describe functions, activities and topics. These terms are put into the thesaurus but are not considered acceptable for use. • Keyword AAA (CV) also includes non-preferred terms. Enter these and tag as non-preferred terms. • Indicate the Equivalence relationship by linking the non-preferred term to the preferred term. Ensure preferred term has a 'USE' tag. • Ensure terms sort into alphabetical order.
8	• Ensure non-preferred terms are entered at the preferred term. • Indicate the equivalence relationship by attaching a 'Use For' tag to the non-preferred term.
9	• Review the listing and ensure no illogical relationships have been created. • Resolve any conflicts that arise between Keyword AAA (CV) terms and the core agency terms you have added. Identifying conflicts is usually only possible for those familiar with local terminology. Table 9 identifies common errors in thesaurus relationships and how to correct them. • Conflicts arise if a non-preferred term in Keyword AAA (CV) is the same as, or very similar to, a term applied to a core function or activity in the business classification scheme. Such conflicts should have been resolved during the development of the business classification scheme. • If the similarity in meaning is confusing, delete any links to the Keyword AAA (CV) term and tag it as a non-preferred term.

Again, the process to construct a merged functions thesaurus does not have to be conducted strictly in line with the process outlined above. The important point to remember is the relationships must be created according to thesaurus rules.

Figure 3 – Main steps in creating a merged functions thesaurus



2.3 Common errors in the relationships established between terms

Table 9 illustrates some of the common relationship errors that occur when creating a functions thesaurus. It aims to help you identify what may go wrong in the process of tagging a term with a specific relationship and how to rectify the problem. Using a thesaurus software application will reduce the likelihood of these errors occurring. The terms used in the table are sourced from Keyword AAA (CV). Further information on software functionality to validate entries into the thesaurus is available in Appendix C – 'Notes on thesaurus software applications', in the section on business rules for thesaurus validation and data entry.

Table 9 – Identifying problems and solutions with thesaurus relationships

Problem	Solution
Broad/Narrow relationship incorrectly applied Production – **NT**: Publications	Publication – **NT**: Production
Equivalence relationship incorrectly applied Equipment & Stores – **USE**: Stores	Stores – **USE**: Equipment & Stores
Term unavailable for titling records A preferred term cannot be applied to records in the business information management system.	The incorrect tag has been assigned to the term, or no tag has been assigned to the term. A 'preferred term' tag must be assigned to enable users to apply the term.
Related relationship applied at different classification levels Policy – **RT**: Business Case (ie, Activity – **RT**: Topic)	Policy – **RT**: Procedures (ie, Activity – RT**:** Activity)
Multiple entries of terms in the thesaurus Guidelines – **USE**: Procedures Procedures – **NT**: Guidelines (Keyword AAA (CV)) Procedures – **NT**: Guidelines (Agency)	Only one of these entries should exist. In these examples, the term has been repeated as it is entered as an equivalence relationship, then with a Keyword AAA (CV) source, then with an agency source.
Terms are related to themselves Procedures – **USE**: Procedures Procedures – **NT**: Procedures	Instructions – **USE**: Procedures Publication – **NT**: Procedures
'Orphans' – terms with no relationships	Preferred terms should have at least one hierarchical relationship

Exercise 2

Proceeding through the steps to construct a thesaurus as outlined above, develop a thesaurus from the following information:

Community Relations (Function)
Ceremonies (Activity)
Guest Lists (Topic)
Community Relations has a non-preferred term of Public Relations
Community Relations is related to the function of Government Relations

3 MANAGING CHANGE TO A FUNCTIONS THESAURUS

Most Commonwealth agencies have a history of gaining and losing functions. Business information systems and thesauruses have always had to cope with this movement. There are general issues to consider to ensure the functions thesaurus evolves with changes to the organisation. This section explains the reasons why a thesaurus may require updating, and outlines the procedures to implement and document the process successfully.

General information on administrative change and the transfer of functions from one agency to another is available in Archives Advice 27 – 'Handling Administrative Change'.[12]

3.1 Reasons to modify a functions thesaurus

Organisations are dynamic entities, so a functions thesaurus must be responsive in order to remain relevant. For example, the language and terminology used in your organisation will evolve. When change occurs to, or within, your organisation, it will need to be reflected in the functions thesaurus.

The importance of being able to update a functions thesaurus cannot be underestimated. Adequate resources and procedures must be allocated to maintain the currency of the functions thesaurus. Otherwise, there is a danger:

> if the thesaurus is permitted to become monolithic and resistant to change, it can actually hinder both indexing and retrieval.[13]

A functions thesaurus will require modifying and updating when:

- a function is inherited from one or more Commonwealth agencies
- a function is relinquished to one or more Commonwealth agencies
- a function is sold outside the Commonwealth
- business processes change and new activities are undertaken
- terminology in the organisation changes.

These matters are examined below. This section also considers the implications of outsourcing of a function even though, technically, a function does not move when outsourcing occurs, as the responsibility is retained by the organisation.

3.1.1 Managing modifications to a functions thesaurus

Changes in the nature of the organisation's functions and activities will have consequences for your business information systems and recordkeeping tools, including a functions thesaurus. The functions thesaurus must be updated to reflect these changes. You will need to make decisions about the terms in the thesaurus and the relationships between terms.

Changes may not occur frequently but they need to be planned and managed through sound business processes. It is critical to ensure that changes to the thesaurus do not compromise the classification processes for existing records. For this reason, it is vital

---

[12] See www.naa.gov.au/recordkeeping/rkpubs/advices/advice27.html

[13] Uri Miller, 'Thesaurus construction: Problems and their roots', *Journal of Information Processing and Management*, 33(4), July 1997, p. 482.

to assign responsibility at a senior level for authorising changes, and to establish a clear process for resolving conflicts between users. Business rules should also be developed that set out the level of documentation required to record changes made to the functions thesaurus.

3.1.2 Documenting changes to the thesaurus

When amendments are made to the functions thesaurus, they need to be documented. A thesaurus software application may have a built-in history log that saves the amendments into a history file, along with 'user id' and 'date'. However, it is more likely that the organisation's information and records manager will document the alterations in a file housed in the business information system.

Other conditions to consider when making changes to the functions thesaurus are:

- permitting only staff with the appropriate knowledge and skills of thesaurus construction to make changes to the content
- restricting the ability to make amendments by issuing 'user permissions' to authorised staff only
- extending 'user permission' to the database administrator, as global changes to the thesaurus may be required.

If hard copies of the thesaurus are made for reference, keep a record of the location of these copies, so that they can be updated simultaneously. If changes are being made frequently, it may be preferable to keep the number of hard copies to a minimum.

3.2 Modifying a functions thesaurus when gaining a function

Through administrative change, an agency may gain a function from another agency. A number of issues should be considered to ensure the terms and relationships applied to the inherited function are compatible with the terms and relationships already established in the functions thesaurus.

3.2.1 The business classification scheme

If the function being inherited is not part of a business classification scheme developed in accordance with the DIRKS methodology, the inheriting agency will need to conduct a business activity analysis to define the function, and to identify the activities and transactions undertaken to support it. Step B of the DIRKS Manual gives detailed advice on undertaking this process.

If the function being inherited was part of a business classification scheme developed in accordance with DIRKS, the agency gaining the function can extend its existing business classification scheme to accommodate the inherited function. These additions can then be added to the functions thesaurus.

3.2.2 Incorporating terms for a new function into the thesaurus

When incorporating terms for the inherited function into the existing functions thesaurus, you will need to consider the following questions:

- Will terms describing the inherited function – and its activities and topics – be in conflict with terms already existing in the organisation's functions thesaurus?
- Will merging the new terms and their respective statuses be done manually?

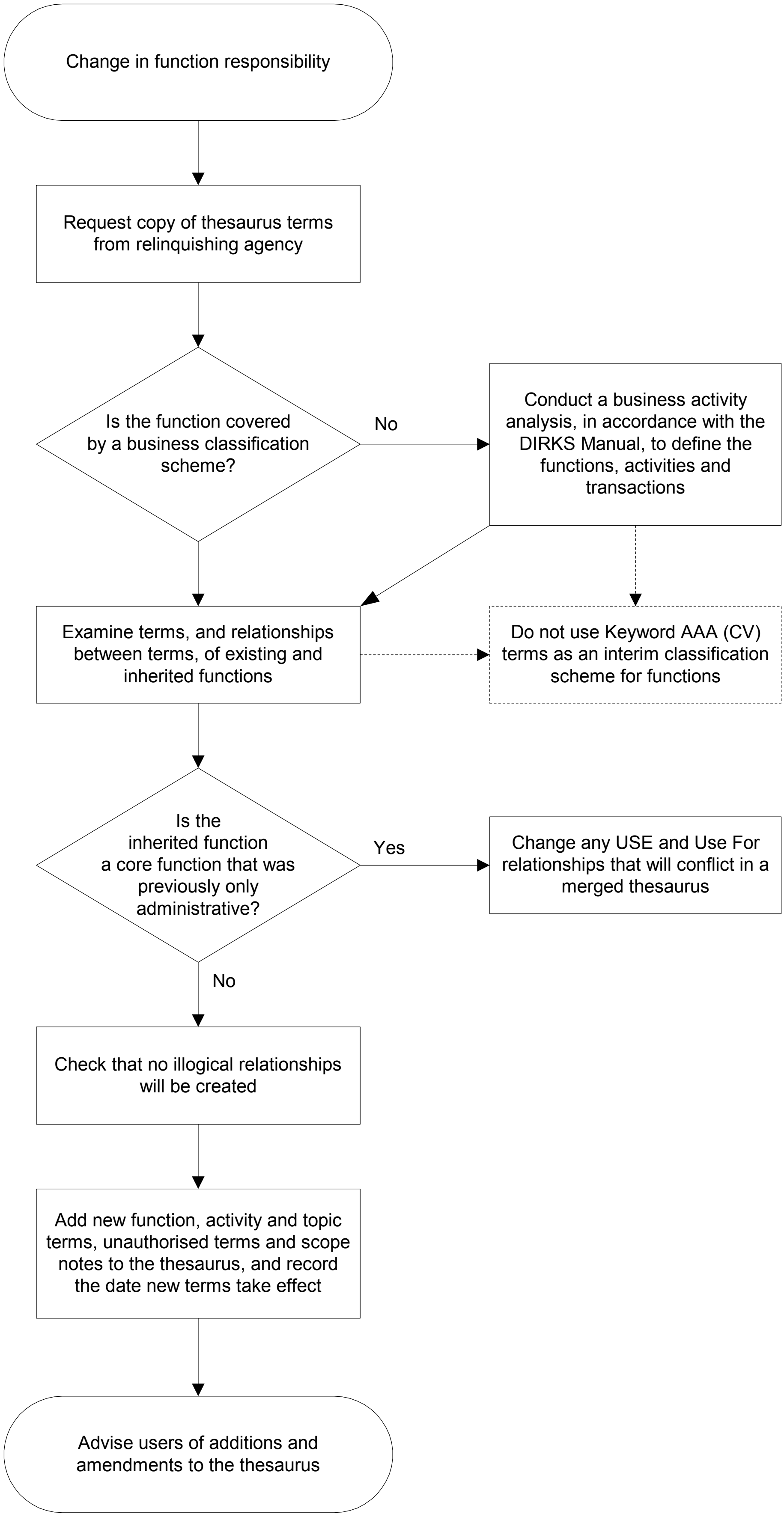
- If the relinquishing agency has created a thesaurus, can you obtain the thesaurus tables in an appropriate format so the terms can be migrated into your thesaurus management software?

Below is an example of how a conflict can arise between terms in the thesaurus.

> Agency A has a core function of 'Publishing'.
>
> Agency B has only an administrative function of 'Publication', from Keyword AAA (CV). In Agency B, 'Publication' has a non-preferred relation with the term 'Publishing' – ie, 'Publishing – **USE**: Publication'.
>
> Agency B inherits the core function of 'Publishing' from Agency A.
>
> Agency B's thesaurus must be amended to remove the non-preferred relationship between 'Publication' and 'Publishing', and to clearly delineate the core and the administrative functions. Both these terms now require a preferred status.

In this example, the changes would have to be done manually. The process required to effect this change is illustrated in Figure 4.

Figure 4 – Tasks to integrate an inherited function into a merged thesaurus



3.3 Modifying a functions thesaurus when losing a function

Through administrative change your organisation may lose a Commonwealth function to another agency, or a function may be privatised. An agency transferring a function must undertake a number of tasks to accurately record the change in responsibility.

A critical issue is to establish the date from which point no new records will be classified or titled against the function that has been lost. The Administrative Arrangements Orders (AAOs)[14] establish the formal change in responsibility. The National Archives posts AAOs on its website.

The actual change may not happen immediately. It is normally up to the agencies concerned to negotiate a handover date. Stopping the creation of new records to match the formal date of the AAO is much easier in an electronic business information system with control over the use of thesaurus terms to describe new records.

When you have agreed on a handover date, record it in the business information system. This date is effectively the closing date for files. The functions thesaurus must be amended to prevent further classification and titling against the terms used for the function being lost.

To do this:

- change the status of the terms relating to the function from preferred to non-preferred, so users cannot inadvertently classify against these terms
- prepare history notes to document the loss of the function and the date the loss became effective.

If an agency loses a function, and the terms relating to that function cannot be used to classify and title records, it is preferable to assign a 'non-preferred' status to these terms. Add an explanation in the history note to account for the term's change in status. It may take a little while for users to change their search patterns.

3.3.1 Modifying a thesaurus when a function is sold to private enterprise

When a function is sold outside the Commonwealth, one of the main decisions, as far as the thesaurus is concerned, is whether or not records of that function remain within the Commonwealth. This decision is important, irrespective of how much of the function is sold.

If records pre-dating the sale are to remain with the Commonwealth, the descriptions of those business functions must remain in the thesaurus. It is not wise to delete a term for a function that has been sold, as it will still be required for retrieval of the records that are retained.

The functions thesaurus must be amended. To do this:

- change the status of the relevant terms for the function that has been sold to non-preferred, to ensure no further records are classified and titled against these terms
- retain these terms in the thesaurus to assist retrieval of previous records
- create a history note to record the date the function was sold and, therefore, the date when relevant terms ceased being valid for classification.

[14] See www.naa.gov.au/recordkeeping/dirks/dirksman/dirks_A13_aao1.html

3.4 Keeping terminology current in a functions thesaurus

The functions reflected in the thesaurus will evolve; but the naming conventions used to describe and classify these functions will change more frequently. As organisational culture shifts, other terms may come into prominence and become preferred terminology.

3.4.1 Reasons to update terms and relationships

There are several reasons why the contents of, and the relationships within, a thesaurus might change. These may include:

- shifts in organisational culture with new terms gaining prominence
- changes in the status of terms
- identification of additional non-preferred terms by users
- recognition, through usage of the thesaurus, that scope notes or history notes need clarification
- the loss or gain of functions through Administrative Arrangement Orders
- emergence of new activities through changes to business processes.

3.4.2 Requests for changes from users

The procedures to enable users to suggest amendments and additions will need to be established by the Information and Records Management section and promoted throughout the organisation. A form, available through the organisation's intranet, is a good way to register requests for changes. The form could also be available in hard copy. An example of such a form is in Appendix A – 'Thesaurus change suggestion template'.

The intranet location can also clearly state reasons why a term can be altered, added or removed. It should also include reasons why changes cannot be made, to avoid repetition of the same, or similar, requests. This location could also communicate any recent amendments to the thesaurus.

Establish procedures to ensure that only staff with relevant expertise and authorisation can decide if the change is warranted. If a suggested change to a term is not appropriate, authorised staff will need to provide users with reasons the amendments could not be made, and record this information.

Below are some reasons why suggested changes may not be appropriate. A thesaurus should not be altered if the suggested change:

- would create an illogical relationship
- would alter the meaning of Keyword AAA (CV) terms
- would change the meaning of a business classification scheme term
- refers to removal of terms for outsourced functions, as an outsourced function is still the organisation's responsibility.

3.4.3 Amending terms in the functions thesaurus

In a functions thesaurus, with multiple hierarchical relationships, it is critical that a change to the status of a term does not affect other relationships in the thesaurus. All affected relationships following a change need to be checked. If a preferred term becomes a non-preferred term, ensure that other non-preferred terms no longer point users to that term.

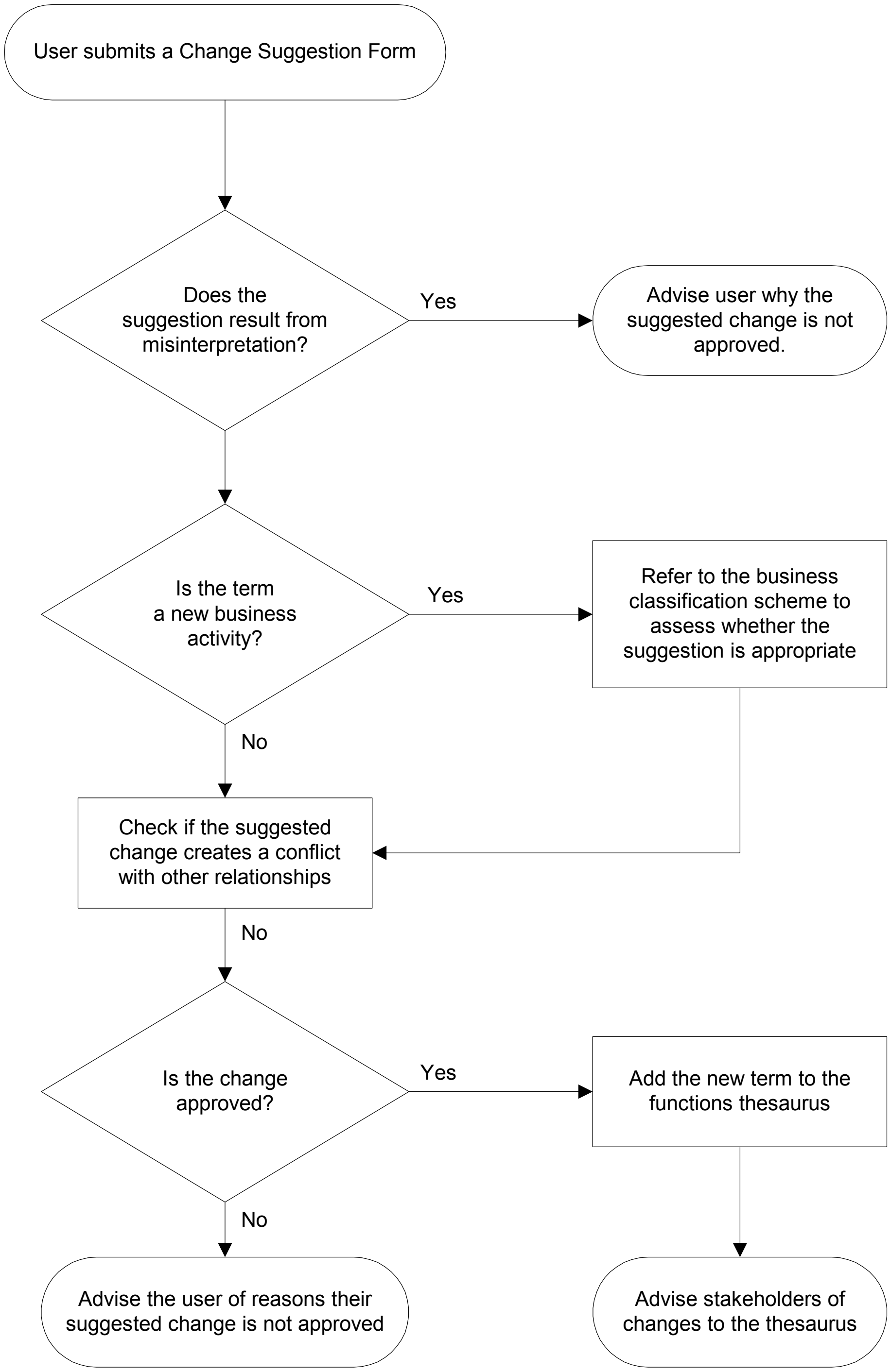
For example, the term 'Legal Opinion' may be superseded by the term 'Legal Counsel'. Previously, 'Legal Opinion' was a valid, preferred term for classifying and titling records and information. Accordingly, this function term is linked to a number of narrower activity terms.

Steps to follow in this scenario include:

- Sever the BT/NT relationships between 'Legal Opinion' and its activities of 'Legal Advice', 'Legal Policy Advice', 'Meetings', 'Reporting', 'Research' and 'Submissions'. This process does not delete any of these terms, only the specific relationship between them.
- Sever the related relationship with the related function of 'Legal Liaison'.
- Sever equivalence relationships (in this case, 'Legal Opinion' had no prior equivalence relationships).
- Tag the term 'Legal Opinion' with a non-preferred status, so it cannot be used to create new records – existing records with the superseded function term 'Legal Opinion' can still be retrieved.
- Add the term 'Legal Counsel' and tag as a preferred term.
- Establish hierarchical BT/NT relationships between 'Legal Counsel' and the activities of 'Legal Advice', 'Legal Policy Advice', 'Meetings', 'Reporting', 'Research' and 'Submissions'.
- Establish related relationships (in this case, 'Legal Counsel' has no related relationships).
- Establish an equivalence relationship between 'Legal Opinion' and 'Legal Counsel' – ie, 'Legal Opinion – **USE**: Legal Counsel'; and 'Legal Counsel – **Use For**: Legal Opinion'.
- Produce a history note to record why 'Legal Opinion' is no longer a valid term for classification and include the date of change of status. Alternatively, this may be recorded in the business information system.

The following flow chart illustrates the main steps for assessing if a suggested change is warranted and then making actual changes to the thesaurus. This applies to new terms, scope notes and existing terms.

Figure 5 – Process to amend terms in the functions thesaurus



3.5 Considerations when a function is outsourced

Outsourcing a government function should not be confused with privatisation or loss of a function to another organisation. When an agency outsources a function, the responsibility for managing the function remains with the agency.

3.5.1 Managing classification and titling with a service provider

Strict business rules need to be developed between the agency and the contracted service provider to ensure consistency in classification and titling. These business rules need to cover the following:

- at a minimum, the service provider must classify and title records in accordance with the agency's classification tools over the life of the contract
- both the service provider's business information system and your organisation's business information system will need to be able to accommodate the outsourced function
- if the service provider uses its own business information system, ensure that it is compatible and that records can be transferred back into the agency's business information system when the contract ends
- advise the service provider of any changes to the status of terms in the functions thesaurus during the term of the contract.

Further general information on outsourcing is available in General Disposal Authority 25.[15]

Exercise 3

Give reasons why the requested amendment can be implemented or why it cannot be implemented.

Term to be reviewed:	**Asbestos**
Suggestion:	Please create a related term relationship between **Asbestos** and **Cases**
Reason for the proposed change:	My section deals with cases of asbestosis
Officer suggesting changes:	Sarah Milne
Position:	Manager
Unit:	Staffing Division

[15] See www.naa.gov.au/recordkeeping/outsourcing/outsource_records/summary.html

APPENDIXES

A Thesaurus change suggestion template

This template is based on the form used by the State Records Authority of New South Wales. It has been slightly modified, with their permission. The template can be used:

- as a record of a request for change and the decision on whether or not the change will be accepted for the agency's merged thesaurus
- as an electronic template with a file number attached for direct lodgement into the agency recordkeeping system
- to suggest changes to Keyword AAA (CV) terms – these suggestions should be referred to the National Archives for further consideration. Agencies should use their discretion as to which suggestions should be forwarded. The Archives may refer various suggestions to the State Records Authority of New South Wales for resolution, if they do not relate specifically to the Commonwealth version of Keyword AAA.

Figure 6 – Suggested template for a thesaurus change form

Term to be reviewed/added: This proposal is for: *(please circle one)* **1.** adding a new term **2.** changing the term's relationships **3.** changing the term's scope note **4.** changing a term to a non-preferred term **The above term could be a:** *(please circle one)* **1.** function **2.** activity **3.** topic **4.** non-preferred term **Broader term:** **Narrower term:** **Related term:** **Preferred terms ('USE' reference):** **Non-preferred terms ('Use For' reference):** **Scope note (include a definition and example of use):**
Reason for the proposed change:
Officer suggesting changes: **Phone no:** **Email:** **Position:** **Unit:** **Date:**
Review by Thesaurus Maintenance Committee (internal) – result: **Approved** o **Not approved** o **Deferred** o **Reason for non-approval:**
Forwarded to National Archives of Australia for further comment: *(yes/no)* **Date:** **Name of officer forwarding:**

B Selection criteria for personnel to develop and maintain a thesaurus

This appendix describes the skills and competencies desirable for personnel engaged in developing a thesaurus. It is divided into four parts:

- recordkeeping competencies
- management competencies
- thesaurus content construction skills
- thesaurus software application competencies.

Recordkeeping competencies

The Recordkeeping Competency Standards[16] issued as part of the vocational Business Services Training Package in 2001 are a valuable source of information on the tasks and products involved in developing a thesaurus and the skills required for undertaking the work. These standards should be used as a point of reference when selecting appropriate officers, either within an organisation or through the contracting process. The standards also give detailed information about the products that arise from the development process.

The most relevant competency standards include:

- **BSBRKG 301A – Control records** – describes the work required to classify, register and track records and maintain information about records within a business or records system.
- **BSBRKG 501A – Determine business or records system specifications** – describes the work required to determine the recordkeeping specifications of a business or records system. Includes ensuring classifications, descriptive terms and metadata are developed into a systems specification appropriate for the selected technology.
- **BSBRKG 530A – Develop and maintain a classification scheme** – describes the work required to develop, or modify, a classification scheme based on an organisation's functions and activities.
- **BSBRKG 504A – Develop terminology for activities and records** – describes the work involved in developing new, or modified, vocabulary controls and in establishing rules for titling.
- **BSBRKG 505A – Document or reconstruct a business or records system** – describes the work required to document or reconstruct a business or records system and its context over a period of time.
- **BSBRKG 603A – Prepare a functional analysis for an organisation** – describes the process of identifying, analysing and defining the boundaries of functions, activities and transactions of an organisation or business unit.
- **BSBRKG 605A – Determine records requirements to document a function** – describes the work involved in determining the structure, content and context of records required to document a function.

[16] Australian National Training Authority, *Business Services Training Package BSB01: Units of Competency for Recordkeeping*, 2001. See www.anta.gov.au for details.

- **BSBRKG 607A – Document and monitor the record-creating context** – describes the work required to monitor and document changes to the particular context or environment in which specific records were or are created, used and maintained. Includes monitoring changes to organisational structures, functions and responsibilities.

Management competencies

- manage a project, develop project plans, and incorporate the needs of stakeholders into the project plan
- conduct interviews with stakeholders, present new concepts, and interpret stakeholder responses
- present oral and written reports on progress with the project
- monitor and explain deviations to the proposed targets
- research and analyse documentation required to support the project
- liaise with a number of stakeholders, some of whom might have conflicting interests
- promote and market the concept and uses of a thesaurus in supporting the business processes of the organisation

Thesaurus content construction skills

- knowledge of the Australian Standard for Records Management, AS ISO 15489 – 2002
- familiarity with the suite of e-permanence recordkeeping standards and guidelines produced by the National Archives of Australia
- understanding of the functional approach to describing business activities and classifying records and ability to:
    - distinguish between a function of an organisation and the subjects that describe the contents of a particular document
    - explain, in a simple manner, the respective roles of a functions thesaurus and a subject thesaurus
- ability to identify changes in the legislative and regulatory responsibilities of the organisation
- ability to analyse the effects of administrative change on the business processes, and how the thesaurus will need to be modified as a result
- ability to develop more than one thesaurus and ensure consistency where appropriate
- ability to integrate terms from different sources into a merged thesaurus
- ability to identify extra terms that are required for the thesaurus
- ability to create appropriate links between terms in the thesaurus according to standards and conventions

- familiarity with the content of Keyword AAA (CV)
- experience in, or knowledge of, linking an agency-specific business classification scheme with Keyword AAA (CV)
- ability to document processes involved in creating and amending a thesaurus

Thesaurus software application competencies

- identify user requirements for a thesaurus
- prepare business rules for a thesaurus software application, and objectives for how the thesaurus might integrate with the other applications that support the business processes of the organisations (such as financial management systems, human resource management systems, database applications, recordkeeping systems) – these business rules are required before a decision can be made on whether to purchase an 'off-the-shelf' package or develop an in-house application
- prepare detailed specifications according to standards and user requirements – these specifications would be required if an in-house application is the preferred option
- prepare a test plan and test potential applications to determine if user requirements are met
- analyse potential applications against user requirements and prepare recommendations for executive approval
- liaise with application development personnel on user requirements for building a thesaurus or incorporating an 'off-the-shelf' package into the business processes of the organisation. Such processes might include:
  - an intranet
  - web-based presentation of agency services and information
  - recordkeeping systems for the classification and retrieval of records

C Notes on thesaurus software applications

This appendix gives general advice on the functionality of thesaurus software applications and outlines some of the questions to ask in order to determine their merits.

Software applications may contribute to the process of creating and maintaining the functions thesaurus as well as assisting users to classify, title and retrieve records and information. These applications may be built into other software, or they might be stand-alone products. A well-designed application will:

- ensure that terms are linked to other terms in appropriate ways – it will not allow inappropriate linkages to be created
- provide online help to aid users in selecting terms for classification
- aid retrieval for the user
- allow browsing by many users concurrently
- identify the source of a term, such as Keyword AAA (CV) or the agency business classification scheme.

Functional requirements for thesaurus software

Thesaurus management tools are available as part of document and/or records management software or as stand-alone products. Stand-alone software can either be used in isolation to the rest of your recordkeeping system or integrated with other software. When planning to implement a thesaurus, the following aspects should be considered:

- Does the software integrate with your records management (and other software) environment?
- Does the software support 'controlled' titling? Is this control determined by a thesaurus, a records classification scheme or both?
- Does the software support as many levels of classification as your agency requires?
- Is the software user-friendly and intuitive?
- Does the technology make records management transparent to the users?
- Can the software create an icon to direct users to browse the thesaurus?
- On leaving the thesaurus software, will the user be taken back to where they started?
- Will the software be used primarily by the end user, or by a limited number of people familiar with thesauruses?
- Does the software enable users to browse the thesaurus to select a term for searching?
- Does the software accommodate special indexing and retrieval requirements?
- In addition to functions-based classification, does the application support a separate ISO 2788 (subject-compliant) thesaurus?
- Have the product developers shown an understanding of records management requirements and retrieval procedures?

- Can the software be linked to the organisation's records disposal authorities and implement access controls, at least at the function and activity levels?
- Can field names and tags denoting the status of relationship between terms be changed to reflect the terminology used in your agency?
- Can the display of results be customised to better reflect the requirements of a workgroup?
- Does the software assign a tag to each term to establish its status?
- Does the software apply a source to each term, in the form of a name or a number?
- Does the software include access controls to ensure only a few officers can actually change the thesaurus?

Business rules for thesaurus validation and data entry

It is important to establish a set of business rules to determine the data entry and validation controls that you want in your thesaurus application. The following conditions are essential:

- a term cannot be linked to itself with any relationship
- the number of classification levels is unlimited, so that a term can always have a narrower term
- a term that has a non-preferred tag (that leads to a preferred term) cannot have a narrower term – but it may have an additional equivalence term
- a term that has a non-preferred tag cannot have a related term tag
- a term is not required to have a hierarchical relationship or a related term – it may only have a scope note
- the same term cannot be linked twice to another term via different tags
- where two terms have a 'related term' association, the relationship only needs to be entered once and the thesaurus will create the reciprocal relationship and display the tag against both terms
- a term cannot be deleted if it has a narrower term tag against it, or if it is associated with a record or other information source description.

Options for users to select terms for classifying records

When users are expected to classify records by selecting up to three terms, and then perhaps adding free text, it is crucial that the mechanism for selecting the thesaurus terms is simple.

Some useful features to include in the software are:

- users may scroll through the valid list, which is in strict alphabetical order
- users may only see the top-level terms in alphabetical order, and then click to see the narrower terms, such as in a directory
- users select the terms from the thesaurus browser and then, behind the scenes, the application populates the appropriate field with the correct combination
- in instances where only one term is required, users can type in the term, as long as it is validated against the list of preferred terms before being accepted.

Using a thesaurus browser application

A thesaurus browser application allows the user to enter a term and have its relationships displayed. Once displayed, there are two ways of applying the correct terms in the business information system. Users can either:

- select an appropriate combination, which will automatically populate the file title field – ie, the browser is linked directly to the data entry application – or
- note the correct combination of terms and then return to the data entry application to enter the record title.

Customising the display of terms

It is possible to customise the way in which terms are displayed in a thesaurus according to the workgroup and workflow processes of users. To customise a thesaurus in this way it is necessary to identify combinations of function and activity terms from the functions thesaurus that most accurately reflect particular business processes, and the records and information they generate.

The agency's application for file creation and titling can then restrict the display of terms to the function and activity terms relating to a particular workgroup.

Users are more likely to be satisfied with a portion of a thesaurus that relates to their function or activity and from which they can more easily:

- identify appropriate terms to classify a record or document
- locate other records relating to their particular functions and activities
- choose a better term than the one with which they started.

For example, users responsible for Staff Development functions and activities would only see the portion of the functions thesaurus relevant to them. Access to the full thesaurus would still be available through the browse application.

Retrieval features

Users should have more than one access point to the thesaurus. A directory style is generally inadequate for users inexperienced with thesaurus relationships or unfamiliar with the terminology, as their starting point for finding terms is not necessarily the top-level terms of the thesaurus.

It is worth investigating applications to see if they can accommodate the following:

- Can a user enter a search term, without knowing whether it is in the thesaurus?
- If the term is a non-preferred term, will it be swapped for a preferred term and the search continued?
- When an activity or transaction is linked to multiple higher-level terms, are the relationships displayed at both levels?

The application should have a variety of display and report options for online and hardcopy output, including the potential for generating *ad hoc* reports. These might include:

- the full thesaurus in alphabetical order
- preferred terms in alphabetical order, with or without scope notes

- preferred terms in hierarchical order, with or without scope notes
- terms from a particular source
- a portion of the thesaurus.

Using a thesaurus in the query and retrieval process

The point at which the thesaurus will be used in the query and retrieval process depends on the nature of the agency's business processes and business information systems. It may be useful to consider the following options.

When a user enters a search term that is in the thesaurus:

- will the term's relationships display before the search is executed, or will all records classified according to that term and all its related terms be retrieved?
- will records with the term in the title be retrieved ahead of records where the term appears in the content of the document?

When a user enters a search term that is not in the thesaurus:

- will the search retrieve records whose free text title includes the term?
- will the search retrieve records whose content includes the term?

D Subject and other thesauruses

An organisation may have, at any one time, multiple thesauruses to describe its records and other business information. The library may have generic subject thesauruses, an in-house thesaurus, or an industry thesaurus, such as those used in the disciplines of education[17] and health.[18] A system that uses multiple thesauruses enables a range of search strategies to maximise retrieval. However, it is important to remember that these other thesauruses do not replace the functions thesaurus. They serve different purposes, such as resource discovery, rather than records classification and retrieval.

Criteria for using a subject thesaurus for classification and retrieval

The value of using subject thesauruses depends on many factors, including:

- the ability of the business information system to handle multiple thesauruses
- the ability to control validation to make sure only preferred terms are applied at appropriate levels – for example, a term from a subject thesaurus can only be used in the file title classification string after a function/activity combination has been entered (this would apply if the third level is a subject descriptor rather than a transaction)
- the extent to which the subject thesaurus has been used previously in the organisation or industry.

Integration of subject thesauruses into retrieval tools

Bearing these factors in mind, there are situations where a subject thesaurus is very useful, for example:

- at the third, or topic, level in the record title – for example, 'Professional relations – Professional liaison – Teachers', where 'Professional relations – Professional liaison' are the Function – Activity part of the string, and 'Teachers' is a term in a subject thesaurus
- to populate the subject element in metadata schemes incorporated in organisational business information systems – the *Recordkeeping Metadata Standard for Commonwealth Agencies* has a subject field in which terms from a subject thesaurus can be entered for further description.[19]
- as non-preferred terms entered into the functions thesaurus

---

[17] The Education Network website uses a variety of education-based thesauruses to help users find specific information and services. See www.edna.edu.au/EdNA – thesauruses include the *Australian Thesaurus of Education Descriptors*, at www.acer.edu.au/library/thesaurus_frameset.html and the *Vocational Education Thesaurus* (VOCED) at www.voced.edu.au/thes.htm

[18] *The Health and Aged Care Thesaurus* is available on the HealthInsite website at www.health.gov.au/thesaurus.htm – you can browse topics or search for a particular topic, and either examine scope notes and related terms or initiate a search for documents with that term.

[19] See www.naa.gov.au/recordkeeping/control/rkms/summary.htm

- when records are published on, and searched for, on an organisation's website or intranet – terms can be mapped to the AGLS Metadata Standard's 'Subject' element, to facilitate web-based discovery[20]
- when records are transferred to the National Archives, and item-level information is recorded in its RecordSearch application – RecordSearch has an optional search field, called 'Keyword headings', which may contain terms from a thesaurus or set of indexed headings.

[20] See the AGLS Metadata Standard at www.naa.gov.au/recordkeeping/gov_online/agls/metadata_element_set.html

E Answers to the Exercises

Exercise 1

Example 1 – True. This set of relationships is represented correctly. The narrower term of 'Litigation' is subordinate to the broader term of 'Legal Services'.

Example 2 – False. This set of relationships is not represented correctly. The representation sets up a cycle where a user will not know which term is the preferred term and which is not. These relationships should be represented by 'Court Cases – **USE**: Litigation'; 'Litigation – **Use For**: Court Cases'.

Example 3 – False. This set of relationships is represented incorrectly. 'Litigation' either has an association relationship or a hierarchical relationship to 'Claims', but not both. This would be properly represented by 'Litigation – **RT**: Claims'; 'Claims – **RT**: Litigation'.

Example 4 – True. This set of relationships is correctly represented.

Exercise 2

Transforming the information supplied produces an alphabetical listing of the terms, beginning with the term 'Ceremonies' and ending with the term 'Public Relations'.

No scope notes have been inserted, as they were not supplied in the information for the exercise. Broader, narrower, association and equivalence relationships have been entered as per the information supplied.

Ceremonies		(preferred term)
BT:	Community Relations	
NT:	Guest Lists	
Community Relations		(preferred term)
NT:	Ceremonies	
RT:	Government Relations	
Use For:	Public Relations	
Government Relations		(preferred term)
RT:	Community Relations	
Guest Lists		(preferred term)
BT:	Ceremonies	
Public Relations		(non-preferred term)
USE:	Community Relations	

Exercise 3

This amendment would not be approved in its current form. The request is to establish an association relationship between 'Asbestos' and 'Cases'. But it is a thesaurus rule that related terms relationships are only established between terms at the same level of classification. In this instance, 'Cases' is a second-level classification term – ie, an activity – and 'Asbestos' is a third- or fourth-level classification term – ie, a topic or subtopic. To address this particular request, the user should be informed that a hierarchical relationship can be established between 'Cases' and 'Asbestos', as the latter is a narrower term.

If 'Asbestos' is not currently listed in the functions thesaurus, and is a valid records classification and titling term, it can be added to the thesaurus and given a preferred status. 'Cases' and 'Asbestos' can be combined together in the titling string in this way, for example:

Human resource management – Cases – Jones, J – Asbestos

Information on extending the use of a classification tool and promoting it to users is available in Part five – 'Maximising the utility of a classification tool' of the *Overview of Classification Tools*.

F Glossary

This glossary defines terms that are central to this document. Most definitions have been extracted from the Australian Standard for Records Management, AS ISO 15489 – 2002, or the DIRKS Manual.

For more comprehensive definitions refer to the Glossary in the DIRKS Manual.[21]

Associative relationship – Covers relationships between pairs of terms which are *not* members of an equivalence set, *nor* can they be organised as a hierarchy in which one term is subordinated to another, yet they are mentally associated to such an extent that the link between them should be made explicit… on the grounds that it would reveal alternative terms which might be used for indexing or retrieval.[22]

Establishes non-hierarchical relationships between related terms.

Authorised term – The terms in a thesaurus that can be applied to a record. Also called 'preferred' terms.

Business function – The largest unit of business activity in an organisation or jurisdiction.[23]

Functions represent the major responsibilities that are *managed* by an organisation to fulfil its goals. They are high-level aggregates of the organisation's activities.[24]

The first level in the business classification scheme of an organisation.[25]

In this document, a function can also be referred to as a core business function, to distinguish it from an administrative function, common to many agencies, such as the functions described in Keyword AAA (CV).

Further information on functions is available in Step B of the DIRKS Manual.

Classification – The systematic identification and arrangement of business activities and/or records into categories according to logically structured conventions, methods, and procedural rules represented in a classification system.[26]

Controlled vocabulary – The use of specified terms and combinations of terms to describe a resource. The opposite of free text.

Equivalence – The relationship between preferred and non-preferred terms where two or more terms are regarded, for indexing purposes, as referring to the same concept.[27]

Relationship linking non-preferred terms to preferred terms.

Function – See *Business function.*

Functions thesaurus – A thesaurus that reflects the unique functions of an agency.

---

[21] See www.naa.gov.au/recordkeeping/dirks/dirksman/glossary.html

[22] ISO 2788, Clause 8.4.

[23] AS ISO 15489, Part 1, Clause 4.15.

[24] DIRKS Manual B.4.2.

[25] AS ISO 15489, Part 2, Clause 4.2.2.2.

[26] AS ISO 15489, Part 1, Clause 3.5.

[27] ISO 2788, Clause 8.2.

A keyword thesaurus produced and maintained by an agency, which has implemented the Keyword Classification System, and which contains keywords, descriptors and forbidden terms. The thesaurus covers terms of a functional nature relating specifically to [an organisation's] specific functions, to provide comprehensive controlled vocabulary over paper and electronic records and recordkeeping systems.[28]

Hierarchy – Relationship based on degrees or levels of superordination and subordination, where the superordinate term represents a class or whole, and subordinate terms refer to its members or parts.[29]

A relationship between terms that is based on a ranking or order from a superior to a subordinate position. For example, the activity 'Child Maintenance Management' is a subset of the broader term 'Legal Case Management'.

Keyword – The highest level term in Keyword AAA. See also *Business function*.

A common data element name in a retrieval interface where the user enters text as a single word, multiple words or a phrase.

Keyword AAA – Keyword AAA is a thesaurus of general terminology designed for use in classifying, titling and indexing administrative records in most technological environments. It covers terminology common to most organisations and should be used in conjunction with a functions thesaurus relating to the organisation's specific functions to provide comprehensive controlled vocabulary coverage. It was developed by the State Records Authority of NSW.[30]

The National Archives has developed a Commonwealth Version of Keyword AAA – Keyword AAA (CV), available to Commonwealth agencies under a whole of government licence.[31]

Keyword classification – Keyword classification involves classifying words into broad, functionally-based areas represented by keywords. These are further defined by descriptors. The keywords and descriptors are the authorised terms provided by the thesaurus. The authorised terms can be supplemented by some free text, that is, words not derived from the thesaurus, such as the name of an organisation, an individual or a project.[32]

Merged thesaurus – A single alphabetical list of terminology for both general and [core] functional terms, which is used to classify records according to the Keyword Classification System.[33] See also *Thesaurus*.

Non-preferred term – The terms in a thesaurus that cannot be applied to a record. Also called 'unauthorised' terms.

---

[28] State Records Authority of New South Wales, *Government Recordkeeping Manual*, Glossary.

[29] ISO 2788, Clause 8.3.

[30] State Records Authority of New South Wales, *Government Recordkeeping Manual*, Recordkeeping in Brief, see www.records.nsw.gov.au/publicsector/rk/aaa/keyword.htm

[31] See www.naa.gov.au/recordkeeping/control/keyaaa/summary.html

[32] State Records Authority of New South Wales, *Government Recordkeeping Manual*, Glossary.

[33] State Records Authority of New South Wales, *Government Recordkeeping Manual*, Glossary.

Preferred term – The terms in a thesaurus that can be applied to a record. Also called 'authorised' terms.

Records classification scheme – A classification system applied to records in a particular business information system, in a particular organisational setting. Hierarchical in nature, it is a system for classifying records based on the business activities that generate the records.

It is a carefully constructed set of terms, business rules and process, which - when applied to records and documents - facilitate their capture, titling, retrieval, maintenance and disposal. It enables records in a business information system to be managed, understood, linked to other related records and retrieved by users.

A records classification scheme should stem from an analysis based on the functions an organisation performs and the resulting business classification scheme.

Scope note – Description to define a term, providing guidance and clarification on the meaning of a term. See Step B of the DIRKS Manual for more detail.

Subject thesaurus – A vocabulary of controlled language for the purposes of describing the content of objects.

Thesaurus – An alphabetical presentation of a controlled list of terms linked together by semantic, hierarchical, associative or equivalence relationships. Such a tool acts as a guide to allocating classification terms to individual records.[34]

In a thesaurus, the meaning of the term is specified and hierarchical relationships to other terms are shown. A thesaurus should provide sufficient entry points to allow users to navigate from terms that are not to be used to the preferred terminology adopted by the organisation. See also *Merged thesaurus*.

Thesaurus application – A software package designed to present words, and their meanings, in a structure that enables the display of relationships according to pre-determined thesaurus structure conventions. The application stores terms and gives them tags and stores their relationships with other terms. It sets validations to prevent non-preferred terms being used for classification and titling.

The software should also have a browse facility to enable users to enter terms. The software then displays relationships and links to search engines to retrieve the records.

Unauthorised terms – The terms in a thesaurus that cannot be applied to a record. Also called 'non-preferred' terms.

[34] AS ISO 15489, Part 2, Clause 4.2.3.2.

G References

Australian National Training Authority, *Business Services Training Package BSB01: Units of Competency for Recordkeeping*, 2001, www.anta.gov.au

Australian Standard for Records Management, AS ISO 15489 – 2002, www.naa.gov.au/recordkeeping/rkpubs/advices/advice58.html

Australian Standard for Records Management, AS 4390 – 1996, www.naa.gov.au/recordkeeping/rkpubs/advices/advice58.html

International Standard ISO 2788 – 1986, *Documentation – Guidelines for the Establishment and Development of Monolingual Thesauri*

Lopez-Huertas, Maria, 'Thesaurus structure design: A conceptual approach for improved interaction', *Journal of Documentation*, 1997.

Miller, Uri, 'Thesaurus construction: Problems and their roots', *Journal of Information Processing and Management*, 33(4), July 1997.

National Archives of Australia, AGLS Metadata Element Set, Version 1.3, 2003, www.naa.gov.au/recordkeeping/gov_online/agls/metadata_element_set.html

National Archives of Australia, 'Archives Advice 27 – Handling administrative change', June 2002, www.naa.gov.au/recordkeeping/rkpubs/advices/advice27.html

National Archives of Australia, *Australian Governments' Interactive Functions Thesaurus*, www.naa.gov.au/recordkeeping/gov_online/agift/summary.html

National Archives of Australia, *DIRKS: A Strategic Approach to Managing Business Information* (DIRKS Manual), September 2001, www.naa.gov.au/recordkeeping/dirks/dirksman/dirks.html

National Archives of Australia, *Keyword AAA: A Thesaurus of General Terms, Commonwealth Version* (Keyword AAA (CV)), May 2001, www.naa.gov.au/recordkeeping/control/keyaaa/summary.html

National Archives of Australia, *Overview of Classification Tools for Records Management*, July 2003, www.naa.gov.au/recordkeeping/control/tools.html

National Archives of Australia, *Recordkeeping Metadata Standard for Commonwealth Agencies*, 1999, www.naa.gov.au/recordkeeping/control/rkms/summary.htm

National Archives of Australia, *Records Issues for Outsourcing including General Disposal Authority 25*, 1998, www.naa.gov.au/recordkeeping/outsourcing/outsource_records/summary.html

Tower Software Pty Ltd, *Advancing Electronic Records Management*, 2000.