

TAFE

Data Processing Protocol

2018-2020

Version 2.1

Copyright Agency Limited

Level 12, 66 Goulburn St

Sydney NSW 2000

ABN 53 001 228 799

www.copyright.com.au

e: info@copyright.com.au

t: 02 9394 7600

f: 02 9394 7601

Revision/Amendment History

EDITION NO.	DATE	SUMMARY OF CHANGES	CONTRIBUTORS
0.1	28/02/17	1 st draft – created from draft DPPTAFE National 2016-17 v0.3 and reformatted	CA
0.2	01/03/17	Inserted Corridor copying Updated Data Quality	CA
0.3	12/07/17	Student numbers under section 3.3 removed, replaced with new paragraph as per DB advice	CA, DB
1.0	28/08/17	All changes agreed and accepted	CA, DB
2.0	13/11/17	References to Distribution Only processing removed Material Type definitions 3.6.6 and 3.7.1 refined Accompanying Artistic Works section deleted	CA
2.1	30/03/20	Title date updated from 16-17 to 18-20 Removed Appendix A (Government Entities List) Updated reference from 'chapter' to 'text' 3.7 and 3.7.1 Removed reference to NCP Corridor Copying 2.1 and NCP section 2.2.3	CA

Contents

TAFE	1
Data Processing Protocol	1
2018-2020.....	1
1 INTRODUCTION	5
1.1 OVERVIEW.....	5
1.2 SCOPE	5
1.3 SURVEY YEAR/PERIOD	6
1.4 CURRENCY.....	6
1.5 CONFIDENTIALITY	6
2 MONITORING SYSTEM	7
2.1 HIGH LEVEL PROCESS SUMMARY	7
2.2 DATA COLLECTION, VALIDATION AND DELIVERY TO COPYRIGHT AGENCY	7
2.2.1 Method 1: Central Print Monitoring System (CPMS).....	7
2.2.2 Method 2: Electronic Use System (EUS).....	9
3 DATA PROCESSING	12
3.1 ROLE OF DATA RESEARCHERS	12
3.2 REFERENCE TOOLS	12
3.3 COPYING VOLUMES	13
3.4 VOLUME CHANGES.....	14
3.5 MATERIAL SOURCE AND TYPE.....	14
3.6 MATERIAL SOURCE	15
3.6.1 Book.....	15
3.6.2 Newspaper.....	15
3.6.3 Journal.....	15
3.6.4 Magazine	16
3.6.5 Website.....	16
3.6.6 Other.....	16
3.7 MATERIAL TYPE.....	17
3.7.1 Text*	17
3.7.2 Article.....	17
3.7.3 Poem	18
3.7.4 Short Story	18
3.7.5 Play.....	18
3.7.6 Artistic Work.....	18
3.7.7 Music	19
3.8 COLLECTIONS OF READINGS / COURSE PACK.....	19

3.9	POWERPOINT.....	19
3.10	COPY TYPE	20
3.11	MORE ON INTERNET MATERIAL	20
3.11.1	Allocations to unauthorised material on websites	20
4	FURTHER RESEARCH AND EXCLUSIONS	21
4.1	OUT OF COPYRIGHT.....	21
4.2	PUBLIC DOMAIN.....	21
4.3	TAFE OWNED WORKS.....	21
4.4	COPYRIGHT AGENCY'S REGISTER OF DIRECT LICENCES	21
4.5	COMMONWEALTH / GOVERNMENT MATERIAL.....	22
4.6	JOINT COPYRIGHT.....	22
4.7	INSTITUTE-SPECIFIC DIRECT LICENSES AND PERMISSIONS.....	22
4.8	OUT OF SCOPE	23
4.9	CLIPART.....	24
5	DATA QUALITY CONTROL	25
5.1	DATA INTEGRITY CHECKS	25
5.2	QUALITY CONTROL.....	25
6	PROCESSING REVIEW	26
	APPENDIX A: OUT OF COPYRIGHT	28
	APPENDIX B: LIST OF ARTISTIC WORKS CATEGORIES	29
	APPENDIX C: VISUAL MATERIAL WHICH IS PROCESSED AS "TEXT ONLY"	30
	APPENDIX D: PROTOCOL FOR THE PROCESSING OF WEBSITE COPYRIGHT NOTICES	31

1 INTRODUCTION

1.1 OVERVIEW

The Copyright Agency (Copyright Agency) and each of the State and Territory jurisdictions represented by the TAFE Copyright Advisory Group (“CAG TAFE”) have an Agreement that covers the monitoring of reproduction and communication of copyright works under Part VB of the Copyright Act 1968 (“the Act”).

Under those Agreements Copyright Agency and CAG TAFE jurisdictions have an agreed data collection system which comprises three data collection methods customised for each of the participating TAFEs. Data obtained through this collection system will form the basis on which Copyright Agency makes distribution of licence fees to rightsholders of copyright works.

The purpose of this document is to establish protocols for processing the type and volume of data collected.

The combined efforts of CAG TAFE and Copyright Agency are aimed at providing integrity, confidence and transparency in the survey processes and outcomes.

This protocol also provides the methods by which rightsholders are identified, for the distribution of payment.

1.2 SCOPE

This Protocol documents the agreed treatment of survey data generated from the TAFE Monitoring System. The Protocol is designed to:

- provide transparency to Copyright Agency and CAG TAFE;
- identify any ‘grey areas’ and establish a platform for consultation and dispute resolution; and
- guide Copyright Agency staff in the consistent and impartial processing of survey data.

The Protocol provides:

- an overview;
- a description of the monitoring system;
- a description of data collection methods and validation processes;
- Copyright Agency’s data setup, processing and exclusions approaches;

- Copyright Agency's Quality Control and CAG TAFE Review Process;
- Copyright Agency's dataset reporting procedure; and
- additional supporting documentation (as appendices).

1.3 SURVEY YEAR/PERIOD

This protocol applies to survey data collected under the TAFE Monitoring System in 2018-20 and may be extended by agreement between CAG TAFE and Copyright Agency.

1.4 CURRENCY

Copyright Agency will ensure this document is kept up to date on an agreed basis to reflect shared stakeholder insights gained during the survey and processing of the data.

1.5 CONFIDENTIALITY

Copyright Agency must keep confidential all details of an Institute's direct licences and only release them to such of Copyright Agency's employees who have a need to know.

2 MONITORING SYSTEM

2.1 HIGH LEVEL PROCESS SUMMARY

An ad hoc data collection system is being implemented, with customisation required for each TAFE, depending on its approach to copying and communicating copyright material. There are three different data collection methods in operation:

Data Collection Method	Areas of TAFE
1. Central Print Monitoring System	CP
2. Electronic Use System	LMS/ LCMS

2.2 DATA COLLECTION, VALIDATION AND DELIVERY TO COPYRIGHT AGENCY

2.2.1 Method 1: Central Print Monitoring System (CPMS)

Summary

The CPMS is implemented in one TAFE which has a Central Print facility (including Central Print facilities shared by multiple institutes and TAFE's which outsource their printing to a third party). IP Copyright Services (IPCS) is facilitating the collection of data from central print facilities, on behalf of CAG TAFE.

Each of the selected TAFE institutes records all copying done by the Central Print facility during the monitoring period in the format required by the Job Spreadsheet.

IPCS delivers in electronic format to Copyright Agency two sets of survey information and data:

1. An MS Excel spreadsheet of jobs metadata; and
2. A set of files captured (named using the JOB NUMBER listed on the spreadsheet) by the relevant institute.

A full copy of every job is provided to Copyright Agency, except those classified as Confidential/Personal/Sensitive or External.

Job Spreadsheet

The Job Spreadsheet contains the following job metadata:

- Date (dd/mm/yy)
- Unique Job No
- Dept Name or code
- Saved File (yes/no)
- No. Of Pages
- No. Of Copies
- Job Type
- Details

Week 1 spreadsheet and files will be sent to Copyright Agency by the end of week 2 of monitoring. Weeks 2 – 10/11 files and W1 – W10/11 spreadsheet (with job status information) will be sent to Copyright Agency by the end of week 17/18.

The original spreadsheet from IPCS is stored without changes and a working copy (the Copyright Agency spreadsheet) is created and saved. Some formatting of the columns is required on the working copy before the spreadsheet can be imported into the Copyright Agency database for technical compatibility only. Original data is not modified in any way.

At some TAFEs the copy of the job may include the original print request form that was sent to the Central Print Facility. Where this occurs the print request form may be included in the volume reported in the “No. of pages” field of the Job Spreadsheet and may need to be subtracted from the total.

Job Type

Before providing the Job Spreadsheet to Copyright Agency, IPCS is responsible for classifying each record according to the agreed list of categories:

- All Institute Materials
- Permission
- Institute Materials and Permission
- Institute Materials and As attributed
- For an external party
- Confidential, Personal or Sensitive
- Part Institute Materials (specify details)
- Part Permission (specify details)

- As attributed
- Source(s) not identified
- Other (specify details)

Where it is necessary to contact the TAFE to verify the Job Type but it has not been possible to establish contact, the record is classified as “Other (specify details)” and “no information available to be provided” is entered in the “Details” field.

Missing Data

Copyright Agency will review the data to perform a quality check – that is, that the required fields and data are compliant with the agreement.

In the case where a field is in error, missing or where insufficient details have been provided (including where digital files cannot be opened or have not been supplied where required), Copyright Agency will raise the issue, within the timeframe set out in the CPMS Institute Management Plan, with IPCS who will investigate the causes. Copyright Agency and IPCS must agree to a course of action and will make reasonable efforts to resolve the issues within the timeframe set out in the CPMS Institute Management Plan.

2.2.2 Method 2: Electronic Use System (EUS)

Summary

Data regarding electronic use of copyright material is obtained from the institute’s Learning Management System (LMS) and/or Learning Content Management System (LCMS). IPCS facilitates the collection of electronic use data, on behalf of CAG TAFE.

IPCS reports digital materials made available online for access by students.

To supply electronic use data:

1. IPCS undertakes a consultation visit with the selected TAFE.
2. Following this visit a report is created, listing all files contained in the LMS/LCMS (note that this is a list of files with some metadata. Attachments are not included at this early stage)
3. The list is examined to exclude files which are clearly out-of-scope, or unlikely to contain remunerable material.
4. If necessary, a sample of the remaining records is selected.
5. IPCS then attempts to extract all (selected) files.

Consultation Visit

At the consultation visit, IPCS will identify the reporting capabilities of each system and the extent and nature of the capability of the Institute to report and provide data. Where an institute makes content available to students using both LCMS and LMS, IPCS will determine whether to report from the LMS, LCMS or both. The reporting systems and capabilities vary for each institute.

Report of files on LMS/LCMS

The selected Institute runs a report (using scripts provided by IPCS) from the LMS/ LCMS, providing a list of digital materials made available online for access by students.

IPCS will identify the nature and extent of the data to be reported/extracted/captured at each TAFE. Any relevant metadata (e.g. intended audience, course enrolment or hits/times accessed) will be provided whenever it is available. As a consequence there will be some variation between institutes in the fields supplied.

An example of the fields reported:

- File ID
- Course ID
- Course URL
- Course Name
- Course Section
- Students per Course
- mimeType
- Category ID
- Category Name
- Users Name
- Username
- Filename
- Source Name
- Date created

Pre-exclusions

Many of the files listed in the report are identifiable from their file name, file type and/or other attributes contained in the report as not containing remunerable material, or unlikely to contain remunerable material.

With the objective of minimising the number of non-remunerable records that require full research, for each dataset provided CA will apply the following process steps prior to IPCS extracting the LMS data files:

- All except a limited list of document types will be pre-excluded. The only document types that will be extracted in the first instance are as follows:
 .pdf, .docx, .doc, .ppt, .pptx, .rtf, pub and potentially some .zip, .msg, .eml and .celtx.
- Files identified as being available to 0 students are excluded.
- Duplicate filenames will be consolidated where they appear to be the same file e.g. same course, same resource ID. Methods of identifying duplicates vary between institutes.

The records will be culled further prior to research by applying pre-exclusions based on key-words. All document file names will be matched against a list of key words and phrases to identify potential TAFE-owned and licensed works. The list will be customised to each TAFE with the assistance of IPCS. CA will review the results of the matching script and confirm which records are to be excluded from the file. Not all documents matched to the key words will necessarily be excluded. The key-word matching process will only be used to provide a list of “possible” exclusions, which will then be confirmed by the CA and IPCS teams.

A full list of the words and phrases used to apply pre-exclusions are provided to IPCS by CA to apply to the EUS spreadsheet.

Sampling

In some cases where the datasets are large, sampling of material will be required. In each case IPCS will consult with Copyright Agency and CAG TAFE on the extent and details of sampling. This may involve a degree of pre-processing in order to identify and exclude whole directories and/or file names that are deemed unlikely to contain copyright material.

Included document types extracted

IPCS will attempt to extract all files listed on the Sample Report. Relevant annotations will be provided, e.g. reasons for document being unavailable.

Licences and Permissions

IPCS will provide CA and CAG TAFE with a list of institute specific licences and permissions. This will be supplied within 4 weeks of the extracted files.

3 DATA PROCESSING

3.1 ROLE OF DATA RESEARCHERS

Copyright Agency employs Data Researchers with expertise and experience in the retrieval of bibliographic information.

It is the role of the Data Researcher to use various resources to:

- verify (records may need to be amended if incorrect);
- validate (no changes required);
- identify copyright owners; and
- complete (add missing information such as author, publisher full title, etc.) information.

In performing these tasks, researchers employ their best judgment based on reliable research tools to assist in reaching decisions about copied material.

It is not always possible for researchers to retrieve comprehensive bibliographic information. Researchers are constrained by the quality of original data provided in the survey and the complex nature of the publishing industry. However, the role of the researcher is to identify and complete the citation as far as possible in the circumstances.

3.2 REFERENCE TOOLS

Copyright Agency researchers will use the following sources of information to identify rights-holders and material copied:

- Copyright Agency's database of works and copyright owners;
 - Copyright Agency membership files, including members' Records of Copyright Material;
 - Previous survey and distribution data;
- Online sources, including:
 - Online Library Catalogues, specifically Libraries Australia
 - TAFE and VET sector websites
 - Ulrichsweb Global Serials Directory
 - Publisher websites
 - Decayed websites – www.archive.org

- WHOis registry

Copyright Agency's processing systems include an audit trail so that all processing activities are recorded. Any alteration to a collected record can be traced to the original source record.

Should insufficient detail prevent a record from being processed (e.g. where digital files can't be opened or haven't been supplied where required), Copyright Agency will query the record with IPCS who will investigate.

3.3 COPYING VOLUMES

The copying volumes are imported into Copyright Agency's Database as recorded by the survey participant on the forms and spreadsheets provided.

Title pages, versos, contents listings and blank pages are not included in the calculation of pages copied.

The total number of pages copied is obtained by multiplying the number of published pages copied by the number of copies of each page. In some situations the researcher may change the number of pages copied after they have counted the number of pages in the job provided (see section 3.4 below).

Any copies made for the specific purpose of attaching to the CA survey spreadsheet will be excluded from volume counts.

For LMS data, IPCS provides course enrolment figures in the LMS Report. Copyright Agency uses these to calculate copying volumes (i.e. number of students who have accessed the relevant file).

3.4 VOLUME CHANGES

The following are examples of situations in which a researcher may change the volume quantity.

- a) Newspaper articles which continue over 2 or more pages are entered as 1 page: *pages revised down*
- b) Where the job provided contradicts the number of pages marked on the spreadsheet or form, changes to the volume may occur. Examples include:
 - reference pages such as title page and ISBN page have been included in count: *pages revised down*
 - survey participant has written 1 page, but 2 pages have been attached: *pages revised up*
 - survey participant has written 1 page, but 2 copyright pages have been photocopied onto 1 side of A4 paper (usually indicated by 2 page numbers or a fold mark down the middle): *pages revised up*
 - unpublished material such as teacher's own or Institute or peak body created material has been included in count along with published material: *pages revised down by the number of unpublished pages*
 - survey participant has written 1 page, but a teacher created worksheet has been attached – each copyright item used is entered separately: *pages revised up or down depending on attachment*
 - survey participant has written 2 pages and attached 2 pages, but they are the same page: *pages revised down*

Note: Any changes made to the number of pages copied, are recorded in the Copyright Agency database.

3.5 MATERIAL SOURCE AND TYPE

Copyright material is classified by researchers into the source from which it originated (“source”) and the type of work that it is (“type”). The Material Source and Type is determined by the researcher using the Guidelines detailed below.

Material Source	Material Type
Book	Artistic Work
Newspaper	Text
Journal	Play
Magazine	Music

Website	Poem
Other material	Article
	Short Story

Material Source is referred to as the Major Work Type in the Copyright Agency Database.

Material Type is referred to as the Minor Work Type in the Copyright Agency Database.

3.6 MATERIAL SOURCE

3.6.1 Book

Material will be classified as a book where:

- there is an ISBN;
- it is listed in a database of books;
- it is listed as a book in a library catalogue (e.g. National Library of Australia); or
- it is a work which is divided into chapters.

3.6.2 Newspaper

Material will be classified as a newspaper where:

- the work is a daily newspaper; or
- the work is a periodic publication in the form of a broadsheet or tabloid, which commonly contains news, comment, features and advertisements.

An ISSN is often an indicator that a work is a newspaper, magazine or journal.

3.6.3 Journal

A Journal is defined as a periodic publication NOT generally available in a newsagency.

Material will be classified as a Journal if:

- it has the word “journal” in its title or subtitle;
- it is principally available by subscription; or
- it contains information of the following nature:
 - technical
 - scholarly

- professional
- academic
- scientific

An ISSN is often an indicator that a work is a newspaper, magazine or journal.

Note: for volume reporting purposes, a Journal is classified as a Periodical (P)

3.6.4 Magazine

A Magazine is defined as a periodical publication which is generally available in a newsagency and published no less than weekly.

Material will be classified as a magazine if:

- it is a periodic publication which contains popular or general interest material, generally published no less frequently than monthly; and
- it is not classified as a journal (J) in accordance with this Protocol.

An ISSN is often an indicator that a work is a newspaper, magazine or journal.

Note: for volume reporting purposes, a Journal is classified as a Periodical (P)

3.6.5 Website

Material will be classified as Website where:

- there is a URL or the work is clearly from a website; and
- the original source is not already classified as Book, Newspaper, Journal or Magazine.

Whenever a webpage is classified as Book, Newspaper, Journal or Magazine, the record will be flagged as being sourced from the Internet.

3.6.6 Other

Copying of the following sources will be classified as Other:

- where publication source details have not been provided but Copyright Agency can identify a creator (the researcher will enter “No publication details provided” on the record)
- CD ROMs (source) – not being one of the above material sources;
- annual Reports;
- pamphlets;

- transcripts, including Hansard;
- examination papers (these should be excluded where they are used for assessment); and
- anything that cannot be classified as one of the above sources.

3.7 MATERIAL TYPE

3.7.1 Text*

Material will be classified as Text if:

- the material is a form of writing other than article, poem, play, or short story as outlined below. Text will have its ordinary meaning, though reference is made to the Macquarie Dictionary definition: 'the main body of matter in a book or manuscript'. Examples include:
 - a page or pages from a book
 - a webpage with text

or

- the pages cannot be classified as one of the other Material Types as described below.

If more than one Text work appears on a page (such as a new chapter by a different author), the first will be processed as the remunerable page and included in the remunerable dataset. If the first Text work is out of copyright, the next Text work in copyright will be processed.

* Due to existing system labelling, material classified as 'Text' appears in the CA database as 'Chapter'.

3.7.2 Article

Material Type will be classified as article if the material is an article or part of an article. Article will have its ordinary meaning, though reference is made to the Macquarie Dictionary definition: 'a piece of writing on a specific topic forming an independent part of the literary publication'.

If more than one article appears on a page, the first article will be processed as Article. This article is taken to be the remunerable article and included in the remunerable dataset. If the first article is out of copyright, the next article in copyright will be processed as Article.

3.7.3 Poem

Material Type will be classified as a Poem if the material is a poem or part of a poem. 'Poem' will have its ordinary meaning, though reference is made to the Macquarie Dictionary definition: 'a composition in verse, especially one characterised by artistic construction and imaginative or elevated thought'.

If more than one poem appears on a page, the first poem will be processed as Poem. This poem is taken to be the remunerable poem and included in the remunerable dataset. If the first poem is out of copyright, the next poem in copyright will be processed as Poem.

3.7.4 Short Story

Material Type will be classified as a Short Story or part if the material is a short story or part of a short story. Short Story will have its ordinary meaning through reference made to the Macquarie Dictionary Definition of Story: 'a narrative, whether true or fictitious, in prose or verse but without metrical structure'.

If more than one short story appears on a page, the first short story will be processed as Short Story. This short story is taken to be the remunerable short story and included in the remunerable dataset. If the first short story is out of copyright, the next short story in copyright will be processed as Short Story.

3.7.5 Play

Material Type will be classified as a Play or part if the material is a play or part of a play. Play will have its ordinary meaning, though reference is made to the Macquarie Dictionary definition: 'a dramatic composition or piece; a drama' (a drama being 'a composition in prose or verse presenting in dialogue or pantomime a story involving conflict or contrast of character, especially one intended to be acted on the stage'.

If more than one play appears on a page, the first play will be processed as Play. This play is taken to be the remunerable play and included in the remunerable dataset. If the first play is out of copyright, the next play in copyright will be processed as Play.

3.7.6 Artistic Work

Material Type will be classified as an artistic work if the work:

- is a map, illustration, photograph, drawing or picture or any other kind of artistic work recognized by the Copyright Act (1968); and
- is a stand-alone artistic work.

If more than one artistic work appears on a page, the first artistic work will be processed as Artistic Work. This artistic work is taken to be the remunerable artwork and included in the remunerable dataset. If the first artistic work is out of copyright, the next artistic work in copyright will be processed as Artistic Work.

All works with a material type of Artistic Work will have an artistic works descriptor. See Annexure D for a list of artistic works categories. See Annexure E for visual material that is processed as text only.

3.7.7 Music

Material will be classified as music if:

- the work is sheet music (including accompanying lyrics);
- the work is song lyrics only (not accompanied by music);
- the work is music from a music tutorial. Music tutorials tend to be books about music or music theory, with instruction and exercises. Tutorials include 'how to' play an instrument (e.g. scales; fingering charts and practice pieces); or
- the person completing the record has specified that they have copied print music and any attachments provided do not suggest otherwise.

Each music record will be given a sub-classification of sheet music, lyrics, or music tutorial.

Record/CD covers (the artwork) are NOT classified as MUSIC.

3.8 COLLECTIONS OF READINGS / COURSE PACK

To record instances where copyright materials form a collection of readings, course pack or similar (to be issued as a set to students or staff) a Course Pack Y/N flag is entered into Copyright Agency's systems. Every job number received is to be checked to determine if course pack "Y" or "N" applies to that record.

The following rules apply in determining a 'Collection of Readings':

- "Yes" will be entered into the database if there is a collection of more than 2 works, consisting of more than 2 consecutive pages each.
- "No" will be entered into the database for all other copying.

This is not an agreed definition between CAG TAFE and Copyright Agency.

3.9 POWERPOINT

A page of PowerPoint slides is counted as one page, regardless of how many slides are reproduced on that page:

- 1 PP slide on 1 page = 1 page
- 6 PP slides on 1 page = 1 page

Where a standalone artistic work (or any other whole work e.g. poem) appears on a slide, it is included in the remunerable dataset.

3.10 COPY TYPE

A copy output type is allocated to each copying record.

The copy type is recorded in the Copyright Agency database according to the following codes:

- “Photocopy” – A photocopy of the material
- “Scan” – The material was scanned
- “Print from Electronic Source” – The material was printed
- ‘Communication’ – The material was communicated online

3.11 MORE ON INTERNET MATERIAL

Where the record indicates a homepage or index page of a search engine has been copied (with no description/minor title information), Copyright Agency will exclude these records under the exclusion term ‘Out of Scope’ unless one of the following conditions applies:

- the record is from Google Earth or Google Maps or;
- the homepage or index page of the search engine features copyright works, such as articles, instructional manuals or artistic works.

All Internet Material will be classified in accordance with the agreed ‘Schools Protocol for the Processing of Website Copyright Notices’.

3.11.1 Allocations to unauthorised material on websites

Unauthorised works which appear on a website will not be processed in accordance with the terms and conditions of that website. The record will be flagged as ‘Copying Not Permitted – Unauthorised Copying’ in the Copyright Agency database.

Copyright Agency will only allocate monies to identified, original author/publisher of the work.

A common example is the unauthorised reproduction of lyrics on a website.

4 FURTHER RESEARCH AND EXCLUSIONS

CAG TAFE and Copyright Agency have agreed on a 'Job Type' classification table. (See Section 2.2.1 – Job Type for detail and table). All works which have not already been excluded because of their 'Job Type' allocation will be subject to a Rights Management check to ascertain whether any exclusions apply as stated below.

Note: The relevant records are excluded from the dataset prior to reporting and distribution to rights-holders.

4.1 OUT OF COPYRIGHT

Works that have been protected by copyright, but for which that protection has expired, are out of copyright. For further guidance in determining whether a work is out of copyright, refer to Annexure B.

4.2 PUBLIC DOMAIN

Works in the public domain will be excluded by Copyright Agency. For practicable purposes Copyright Agency processes all national flags, national emblems and road and safety signs as Public Domain works. (Note that non-national flags, e.g. The Olympic, Aboriginal and Torres Strait Islander flags are still in copyright).

Works created by the United States Federal government are not entitled to domestic copyright protection under U.S. law and as such are commonly referred to as being in the public domain. Copyright Agency will classify these works as an exclusion. Many publications by the U.S. Federal government do, however, contain protected works authored by others and these works will be included.

4.3 TAFE OWNED WORKS

TAFE Owned works will be processed as Out of Scope – see 4.8 below.

For usage datasets where a Job Type is provided, these are identified on the spreadsheet as "Institute Materials".

Where any TAFE-owned works contain embedded third party works, other than insubstantial copying, the relevant page containing the third-party work is included in the remunerable dataset for both volume and distribution.

4.4 COPYRIGHT AGENCY'S REGISTER OF DIRECT LICENCES

Where a copyright owner advises Copyright Agency that their works are wholly owned and TAFEs (or more generally, educational institutions) may copy their works for free these works are excluded in one of two ways:

- a) Where the owner provides work titles, the title has an exclusion applied to the work in the Copyright Agency database.
- b) The owner notifies Copyright Agency that their works are free, but is unable to provide a list of titles. An exclusion is applied to the owner's identity in the Copyright Agency database and all works identified are subsequently excluded.

4.5 COMMONWEALTH / GOVERNMENT MATERIAL

All Commonwealth, State and Local Government Australian Material will be excluded other than those works and types of works advised by departments and any other works advised from time to time.

4.6 JOINT COPYRIGHT

This is relevant to newly identified material where there are multiple publishers. If the publisher is a Commonwealth, State or Local Government entity, or it has advised Copyright Agency that TAFE institutes may copy their works for free and the other identified publisher has NOT advised Copyright Agency that their works are freely reproducible, Copyright Agency (Rights Management) will contact all identified publishers to confirm rights control over the copyright material.

Where all the publishers confirm that full or part rights control of the copyright material sits with a publisher who is willing and entitled to collect a Copyright Agency payment, the record will be included.

Copyright Agency will flag all works for which there are joint copyright owners for review by the Processing Review Committee.

4.7 INSTITUTE-SPECIFIC DIRECT LICENSES AND PERMISSIONS

As part of the dataset supplied by CAG TAFE to Copyright Agency under the Monitoring System, participating Institutes notify Copyright Agency of material that has been reproduced or communicated under direct license or where the institute has permission from the rights-holder to reproduce.

This advice is in the form of a Job Type classification of "Permission".

In addition, Copyright Agency will exclude website content in accordance with Appendix D: Protocol for the Processing of Website Notices.

Copyright Notices on Websites

Some websites grant licences or permit copying and/or communication of works contained on the website. These licenses vary greatly in their terms and application. Copyright Agency

researchers will make reasonable efforts to research and establish the terms of website licences.

All material classified as “Internet Material” will be classified according to the Protocol for Processing Website Copyright Notices in Appendix D.

Where Copyright Agency has received direct permission from an Institute or the copyright owner, to copy and/or communicate the material (for example, it is excluded Commonwealth material or subject to a direct licence) this advice will override the website copyright notice.

4.8 OUT OF SCOPE

Material classified as ‘Out of Scope’ will be flagged as excluded work. Below is a list of examples of Out of Scope:

- a) Survey Manager, Copyright Agency or CAG TAFE material relating to the survey or administration of the survey
- b) Personal Material of a survey participant
- c) Material that appears to be for internal administrative purposes (e.g. letters, invoices, student reports, newsletters, timetables etc.). Note: 3rd party embedded works identified within this material is also Out of Scope.
- d) Any copying of quotations and extractions of three (3) paragraphs or less, other than whole works
- e) Material created by the surveyed institution’s state peak body where it has not already been classified as a NEALS non-remunerable work
- f) Material created by any of the state peak bodies or institutions
- g) Material created by the surveyed institution’s current employees –teacher’s own work
- h) Media or press releases
- i) Stand-alone **unembellished** geometric shapes (e.g. Mathematical triangles, squares, circles, angles) and any copying of dice, dominos, blocks, unembellished mazes, playing cards and clock faces **and** there is three (3) paragraphs or less of accompanying text.
- j) Blank tables and blank grids
- k) Where the record indicates a homepage or index page of a search engine has been copied (with no description/minor title information)
- l) Examination papers/materials if used for assessment purposes. If material is dated the same as the survey year, the material will be excluded. (Note: If the exam is dated from a previous year e.g. 2010 or 2007, the material is to be included)
- m) Logos (including app icons on smart-phones, iPads etc.)

- n) Advertisements (including sales catalogues)
- o) Computer generated graphs/patterns, ECG's and seismographs

4.9 CLIPART

Clipart are images whose usage is governed by the terms of individual copyrights and usage rights. Generally these can be categorised by one of the following:

- royalty free;
- CD/DVD Rom licence;
- subscription licence; or
- software (e.g. Microsoft, Apple) licence.

The parties have agreed to assume that the usage of clipart by institutes in jurisdictions represented by CAG TAFE is under licence and all will be excluded accordingly.

When a website has built a library of images in a sub-folder this will not be considered clipart. A researcher will however double-check the terms and conditions of the website to confirm that these images are not royalty free.

5 DATA QUALITY CONTROL

5.1 DATA INTEGRITY CHECKS

During processing and at the end of processing various data integrity checks are carried out by Copyright Agency to ensure that the data has been entered correctly and consistently. This is a separate process to the data validation that happens before point of research.

The purpose of these checks can be broadly characterised as follows:

- identify any records that should be excluded from the data set for both volume estimates and for distribution;
- ensure works of the same type have been entered consistently;
- improve the accuracy and consistency of bibliographic data;
- minimize unidentified records; and
- ensure records are allocated to members consistently.

5.2 QUALITY CONTROL

To verify that the data researchers have processed survey data in accordance with this protocol, Copyright Agency applies an internal quality control check to ensure that:

- No records have been duplicated or omitted. This is done by comparing the entry lines in Copyright Agency's database to the original records supplied.
- What has been entered in Copyright Agency's database is representative of the records provided. This is done by comparing the fields on the records to the corresponding fields in the database, including any amendments made by data researchers.

Any discrepancies identified by the data checker are referred back to the data researcher for clarification and/or correction.

As a final check, system query reports are run across the whole data set to ensure that all data is intact and that there are no invalid entries.

6 PROCESSING REVIEW

It is important that there is transparency in processing so that both Copyright Agency and CAG TAFE can be assured that:

1. the Protocol is being followed; and
2. where uncertainty arises in relation to processing issues, both parties have the opportunity to participate in the determination of the outcome.

In order to meet the requirement in paragraph 1, the following procedures will be adopted:

- a) Each party will appoint up to two people to sit on a Processing Review Committee.
- b) The Processing Review Committee will have broad oversight of the processing operations in so far as they relate to the processing of records for the purposes of assessing remunerable copying.
- c) Copyright Agency will give a quarterly report to the Processing Review Committee which will include details of any issues or uncertainties which have arisen in the preceding quarter and any other matter the Processing Review Committee reasonably requires.
- d) The Processing Review Committee will attempt to resolve any issues or uncertainties. If they fail to reach agreement, they will consider whether it would be appropriate to engage in a dispute resolution process involving a third party.
- e) The CAG TAFE representatives shall be given reasonable access from time to time to view the processing operations undertaken by Copyright Agency and to have those operations explained.
- f) The Processing Review Committee from time to time may recommend amendments to the Protocol to reflect changing circumstances and needs.

To demonstrate that the data has been processed in accordance with this Protocol CAG TAFE may appoint a representative to visit Copyright Agency once a year (to a pre-agreed schedule). This visit will provide an opportunity for the CAG TAFE representative to observe and discuss Copyright Agency's processing procedures and identify any processing issues that have been encountered during research.

One way Copyright Agency may assist would be to provide a random selection of no more than 100 records for CAG TAFE to review (annually). Copyright Agency will then retrieve the records and show all actions taken from when the data arrived to its current state.

Copyright Agency is responsible for the following:

- a) ensuring records provided by the Survey Manager are processed in accordance with this Protocol;
- b) preparing a report each quarter on the processing carried out in relation to survey records and recording any issues or concerns that may have arisen; and

- c) seeking advice and direction from the Processing Review Committee in relation to how records should be treated for the purposes of the surveys.

APPENDIX A: OUT OF COPYRIGHT

	Type of Material	Factors affecting duration	Copyright has expired IF:	Duration if work was still in copyright on 1 January 2005
A	Works first published anonymously or under a pseudonym ¹ (except B)		first published before 1 January 1955	year first published + 70 years
B	Photos first published anonymously or under a pseudonym		taken before 1 January 1955	year first published + 70 years
C	Photographs (except B)		taken before 1 January 1955	life of creator + 70 years
D	Engravings (except A)	published during creator's life	creator died before 1 January 1955	life of creator + 70 years
		not published during creator's life	first published before 1 January 1955	year first published + 70 years
E	Artistic works (except A, B, C & D)		creator died before 1 January 1955	life of creator + 70 years
F	Computer programs (except A)		creator died before 1 January 1955	life of creator + 70 years
G	Literary, dramatic and musical works (except A & F)	made public ² during creator's life	creator died before 1 January 1955	life of creator + 70 years
		NOT made public during creator's life	made public before 1 January 1955	year made public + 70 years
H	Cinematograph films (underlying works such as screenplays and music are also protected)	made before 1 May 1969 and regarded as "dramatic work" ³	creator of film as dramatic work died before 1 January 1955	life of creator of film as dramatic work or of cinematographer (whichever longer)
		made before 1 May 1969 and not regarded as "dramatic work"	made before 1 January 1955	life of cinematographer + 70 years
		made on or after 1 May 1969	none in public domain by 1 January 2005	years first published + 70 years
I	Sound recordings (Underlying works are also protected)		made before 1 January 1955	years first published + 70 years
J	Broadcasts	made before 1 May 1969	not protected at all	not protected at all
		made on or after 1 May 1969	none in public domain by 1 January 2005	Year made + 50 years

¹Where the identity of the author is not known and cannot be ascertained by reasonable inquiry

²"Made Public" here means published, performed, broadcast or recorded and offered for sale

³Films made before 1 May 1969 are regarded as "dramatic works" "the arrangement, the acting form or the combination if incidents represented gives the work an original character"

APPENDIX B: LIST OF ARTISTIC WORKS CATEGORIES

Works which ARE classified as artistic works

- Illustration e.g.
 - drawings
 - etchings
 - diagrams
 - dot-to-dots and colour-by-numbers
 - woodcuts
 - engravings
 - sketches
- Photo
 - film stills
 - photographs
 - photomontage
 - composite photos
- Painting
- Map
- Comic Strip
- Cartoon
- Plan e.g.
 - blueprints
 - architectural plans
 - engineering plans
 - circuitry diagrams

APPENDIX C: VISUAL MATERIAL WHICH IS PROCESSED AS “TEXT ONLY”

- Crosswords
- Tables
- Find-a-words
- Sudoku
- Grids (empty grids now processed as out of scope)
- Graph
- Molecular and chemical structures
- Screen captures without visual material, including web pages
- Languages which are written using characters (e.g. Chinese and Japanese), script (e.g. Cyrillic) or pictorial representations (e.g. hieroglyphics)
- Journalist biography photos
- Text only posters
- Text only certificates
- Geometric shapes which form patterns, or include embellishments such as trees or grass
- Flowcharts

Visual Material processed as “Accompanying Artistic Works”

- Certificates with visual material
- Posters with visual material

APPENDIX D: PROTOCOL FOR THE PROCESSING OF WEBSITE COPYRIGHT NOTICES



2019 Website
Processing Protocol F