
Data Processing Protocol
Agreed between Universities Australia and Copyright Agency
August 2010

University Processing Protocol

A THE APPROACH

1. Introduction

- 1.1 This document sets out the rules for processing of data received from Universities under the Monitoring System.
- 1.2 The procedures set out below generally occur in the sequence outlined, however on occasion, due to resources and timing, they may not occur in the specific order.

2. Role of Data Researchers

- 2.1 CAL employs librarians with expertise and experience in the retrieval of information as data researchers. It is the role of the data researcher to utilise various resources to:

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

information provided by the Universities in accordance with this protocol.

- 2.2 The researcher is also required to identify copyright owners. In performing these tasks, researchers employ their best judgements based on accepted research tools to reach decisions about copied material.
- 2.3 To perform these tasks the researchers utilise the most recent editions of the following;
 - (a) CD Roms such as *Global Books in Print*, *BookBank and Thorpe Rom*, *Australian Compact World*, *Books Out of Print* and *Ulrich's International Periodicals Directory*;
 - (b) CAL's database of works and copyright owners;
 - (c) The Internet;
 - (d) Online Library Catalogues
 - (e) *The National Guide to Government*;
 - (f) Margaret Gee's *Australian Media Guide*;
 - (g) *Directory of Australian Associations*;

-
- (h) *Directory of Acronyms and Abbreviations;*
 - (i) *B& T Advertising, Marketing and Media Year Book;*
 - (j) Press Radio and TV Guides
 - (k) Telephone directories
 - (l) CAL membership files, including members' Records of Copyright Material;
 - (m) Previous survey and distribution data;
 - (n) Publisher directories and catalogues;
 - (o) *The Government Who's Who of Australia;* and
 - (p) *Periodicals in Print.*
 - (q) *CAL's register of excluded works*

2.4 It is not always possible for researchers to retrieve comprehensive bibliographic information and verify rightsholders. Data researchers are constrained by the quality of original data provided under the Monitoring System, as well as the complex nature of the publishing industry. However, the brief is to identify and complete the citation as far as possible in the circumstances.

2.5 The rules set out in this document must remain relevant and up to date. Amendments and adaptations may be required from time to time to reflect and respond to any new issues arising from the data collected or the Monitoring System. CAL should notify Universities Australia electronically and by hardcopy of any proposed amendment or adaptation before its adoption. If within three clear working days after receipt by Universities Australia of such notification the universities do not object to the proposed amendment or adaptation, it may be adopted. If objection is taken, the amendment or adaptation should be referred to the Tribunal to determine whether it should be implemented prior to its adoption.

2.6 So that all data processing activities can be clearly tracked, any alteration to or annotation to data received from Monitoring Body must be clearly identifiable.

B. THE HARDCOPY SYSTEM

3. Initial Data Sorting and Set Up

3.1 CAL receives the Monitoring System record forms periodically in batches – after collection from the institution by Monitoring Body. The record forms are collated and attached to a Record Form coversheet. With the batch of records Monitoring Body include a checklist which details:

- the date the records were dispatched from Monitoring Body;
- a list of the department Ids (as assigned by Monitoring Body);
- a list of the department names;
- number of monitored copiers per department and id; and
- any additional comments relating to the records such as “Copier 2 records held back”.

3.2 The checklist is stamped with the date the records were received at CAL.

-
- 3.3 The copying identity, that is, the particular university and the departments monitored are entered onto the CAL database. The following steps occur:
- The CAL licence number which identifies the university monitored is retrieved from the CAL database;
 - The year of monitoring is added to the profile of the CAL licence number;
 - Each department monitored is allocated a CAL organisation number; and
 - CAL organisation numbers are linked to the Monitoring Body department ID, the year in which the distribution relating to the records will occur, and the distribution pool (that is, “copied for” code).

3.4 Copying records are then sorted by:

- (a) University;
- (b) Department ID;
- (c) copier number; and
- (d) date (period ending).

In the event that no published material was copied in a bundle of records “Nil” is written in red on the record forms.

4. Identification of University Departments Monitored

4.1 The bundles of Published Material Record Forms (“PMR Form”) and the Copying Declaration Forms (“CD Form”) are checked to determine if copying by more than one department appears on the forms. (This usually occurs when copying machines are shared by two or more departments). This determination is based on the entry in question 2 and 3. Where necessary, the name entered in question 3 is searched against the university website.

4.2 If two or more departments are sharing the copier (as determined by the process in 4.1) and all departments are within the survey, CAL will create separate bundles of records for copying by each department.

4.3 If the information provided on Monitoring Body’s checklist is not sufficient to identify the department (e.g an acronym appears) or does not correspond to the name of the department completed on the record form:

- (i) The university web site is searched to determine if the acronym on the checklist can be used to identify the department in full;
- (ii) If so, the process as set out in 4.1, and 4.2 are followed.

4.4 If the name of the department cannot be determined using these means, it is referred to Monitoring Body for resolution.

5. PMR and CD Forms

These forms are checked to determine if question 10 of the PMR form and question 11 of the CD form has “Yes” or “No” ticked:

If there is a Y (Yes) a check is made to determine whether an attachment has been provided. If no attachment has been provided it will be referred to Monitoring Body for resolution; or

Where there is a N (No) no further checks are necessary.

6. Lot Identification

-
- 6.1 A green identification covering sheet is created for each bundle of copying records which includes the following information:
- University name;
 - University Licence Number;
 - CAL's Organisation Id;
 - Monitoring Body Department Id;
 - Monitoring Body Copier Number; and
 - Date (period ending).
- 6.2 All Copying Declaration Forms are identified by writing "copier 99" on each covering sheet.
- 6.3 Lot Ids which provide a unique identifier for each bundle of records are created by entering in the CAL database all of the information as set out in 6.1.
- 6.4 Excel spreadsheets are created by university, by department and by department Id. See attached example for Griffith University.
- 7. Data Processing**
- 7.1 Record bundles are assigned to data researchers. The process outlined in sections 8 to 12 takes place in respect of the copying of copyright material recorded on the forms.
- 7.2 Should insufficient detail be included to allow identification of the copyright owner, CAL will query the record with Monitoring Body to enable follow-up of the staff member concerned, through the Copyright Monitoring System Manager.
- 8. Copying Volumes and details**
- 8.1 The correct copying volumes are entered for each record. Any changes made in accordance with this section 8 are marked in red pen on the original record form.
- 8.2 When processing the PMR Form if there is a discrepancy between the volume recorded in question 4 and the bibliographic details (which show the page numbers from the original work copied), question 4 is amended to be consistent.
- 8.3 Question 6 on the PMR Form shows the total pages copied. It is obtained by multiplying the entry in question 4 (number of published pages copied) by the entry in question 5 (number of times each page copied). If the figure in question 6 does not equal the multiples of 4 and 5, data researchers will combine the information provided in columns 4, 5 and 6 to determine the correct entry for question 6.
- 8.4 In the event that a copy of all of the material has been provided the data researcher will use the actual copy to determine the number of pages of original material that were copied (regardless of the volume recorded on the form attached to the copy) in respect of the PMR and CD forms.
- 8.5 The CD Form is processed in the same way as the PMR Form.

8.6 In the case of high volume copying recorded on the PMR Form (total pages copied is greater than 500) researchers will refer the citation to the Data Manager. The Data Manager will then determine via Monitoring Body whether further verification is required from the university.

8.7 If the total pages copied exceed 10,000 pages on the C D Form, this will also be referred to the Data Manager to determine via Monitoring Body, whether further verification is required from the university.

9. Material Type

The correct material type is allocated to each copying record in accordance with this section 9. Changes to the coding are marked in red pen on the original record.

The material type definitions are as follows;

9.1 Music (M)

Material will be classified as music:

- (i) if the work is sheet music.

9.2 Artistic Work (A)

Material will be classified as an artistic work:

- (i) If the work is a map, illustration, drawing or picture or other kind of artistic work recognised by the Copyright Act 1968 (the Act).
- (ii) If it has been coded as such on the original record form by the person doing the copying

9.3 Legislation (L)

Material will be classified as legislation:

- (i) If the work is an Act passed by the Federal Parliament, State Parliaments and the Legislative Assemblies of the Northern Territory, the Australian Capital Territory and Norfolk Island. Legislation also includes delegated or subordinate legislation made under such Acts.

9.4 Law Reports (LR)

Material will be classified as a law report:

- (i) If the work is a published account of a legal proceeding, which includes a judgment and often other editorial notes.

9.5 Newspaper (N)

Material will be classified as a newspaper:

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

9.6 Journal (J)

Material will be classified as a journal:

- (i) Where the work is a periodic publication that is other than a newspaper.

9.7 Book (B)

Material will be classified as a book:

- (i) Whenever there is an ISBN; or
- (ii) Whenever it is listed in a recognised database of books (see items listed in paragraph 2.3); or
- (iii) Whenever it is listed as a book in a recognised library catalogue (e.g. National Library of Australia); or
- (iv) Whenever it is a music tutorial (a book containing text as well as music, for example a theory or musicianship book)
- (v) Whenever it is a work is divided into chapters.

9.8 Unreported Judgment (U)

Material will be classified as a unreported judgment:

- (i) Where the work is a transcript of a court proceeding (usually a judgment) which is not normally officially reported.

9.9 Web Pages from the Net

One of M, A, L, LR, N, J, B or U dependent on the original source

Anything else that cannot be classified as M, A, L, LR, N, J, B or U will be classified as W

9.10 CD Roms

One of M, A, L, LR, N, J, B or U dependent on the original source

9.11 Other (O)

Copying of the following material will be classified as other:

- (i) CD Roms – not being one of M, A, L, LR, N, J, B or U;
- (ii) Annual Reports
- (iii) Pamphlets
- (iv) Transcripts, including Hansard; and
- (v) Anything else that cannot be classified as M, A, L, LR, N, J, B or U

9.12

Where the university ticks "Yes" in Question 10 on the PMR Form or Question 11 on the CD Form (incidental Artworks), and no copies have been provided CAL will query the record with the Monitoring Body in accordance with section 7.2. Where a copy of the material has been provided, information relating to the artistic works will be entered into the CAL database using the ZM field.

9.13 Materials with accompanying artistic works will be separately identified without any effect on the volumes of copying. For instance a journal article containing 5 original pages with 2 incidental diagrams copied for 10 students would be entered as 5 pages x 10 with material type of "J" and ZM of F, Y or N as follows:

-
- F YES accompanying artistic works with a full FILE copy attached
- Y YES accompanying artistic works but no copy provided
- N Text Only, NO accompanying artistic works
- 9.14 Text only records are entered into the CAL database field as “copy type” “P” Photocopy.
- 9.15 If slides are recorded on the C D Form (question 9), “S” – Slide is entered in the “copy type” field on the CAL database.
- 10. Copied for Code**
- 10.1 The “copied for” data is entered in the database according to the following codes;
- For Distribution to Internal Students (I) = 002
 - For Distribution to External Students (E) = 003
 - For Distribution to Staff (S) = 007
 - For all other purposes (O) = 005
- 10.2 There is no processing of bibliographic data for a single copy made with the copy code “S”.
- 10.5 In the event that two codes have been provided and the volume in question 5 is 2, the data researcher enters two separate instances of copying.
- 10.6 In the event that two or more codes have been provided, and the volume in question 5 is greater than 2, and one of the codes includes the staff (S) code, one copy will be coded ‘S’, and the remaining copies will be coded in accordance with paragraphs 10.1
- 10.7 If question 8 is not ticked, contains a nonsensical entry or is ambiguous the researcher refers to the surrounding records for possible clarification. For instance, a pattern of copying for that person may already be established. When in doubt, the matter is referred to Monitoring Body or the university as per section 7.2 for resolution.
- 10.8 When copies are coded ‘O’, and specifics have been provided, the following sets out a basis for treating and coding copying as VB copying or not, without having all the information to make a more definitive decision, and how the copying processed as ‘VB’ copying will be coded. The ‘non-VB’ copying is excluded from processing altogether and does not form part of the data file. This section 10.8 also applies to the 2002-2003 Monitoring System and for any universities monitored subsequent to the 2002-2003 Monitoring System but prior to the date on which CAL, Universities Australia, and the Universities agreed the changes to this protocol;

VB Copying'Non-VB' Copying

	<u>Coded as to:</u>	(all excluded)
➤ Researchers	S	➤ DD services
➤ Staff	S	➤ Off shore students
➤ Govt depts/Authorities	I	➤ Community groups
➤ Other unis	I	➤ General public
➤ Colleagues	S	➤ Non-uni students
➤ Conference presentation	S	➤ Non-tertiary students
➤ Clients, patients	I	➤ Non-uni bands/orchestras
➤ Clinic clients	I	➤ Parents
➤ External academics	S	➤ Overseas academics
➤ Grant application	S	➤ Overseas requests
➤ Industry researchers	S	➤ School teachers
➤ Lecture overheads	I	
➤ Myself	S	
➤ Readers/Reading graph	I	
➤ Investigation	S	
➤ Professional training	S	
➤ Self	S	
➤ Supervisor	S	
➤ Student noticeboard	I	
➤ Staff noticeboard	S	
➤ Reviewers	S	
➤ Seminar	S	
➤ Visiting academic/Scholar	S	

As this list is not exhaustive, it will need to be updated from time to time. Where the list does not resolve which code to use, the copying will be entered with the 'O' Copied For code

11. Copy Output Type

11.1 The purpose of this step of the processing is to ensure that a copy output type is allocated to each copying record. Any changes are marked in red pen on the original record in accordance with this section 11.

11.2 The output type is recorded in question 9 on the CD Form.

12. Course Packs

12.1 When processing the PMR Form or the CD Form, in respect of question 8;

- (i) If “Yes” has been ticked, the record is entered as “Course Pack Yes”
- (ii) If “No” has been ticked, the record is entered as “Course Pack No”

12.2 In circumstances, where the material copied has been attached to the PMR or CD form and it appears that the copies form part of a collection of readings or course pack or a similar collection, and “No” has been ticked it will be referred to Monitoring Body for resolution.

12.3 If any question has not been completed, or contains a nonsensical code or is ambiguous, the researcher refers to the surrounding records for possible clarification. For instance, a pattern of copying for that person may already be established. In all other cases, the record will be referred to Monitoring Body for clarification by the institution for final resolution by Monitoring Body.

C. THE ELECTRONIC USE SYSTEM

13. Initial Data Set-Up

13.1 CAL receives the monitoring system spreadsheets from Monitoring Body, who will provide covering information such as

- The number of spreadsheets provided
- The names / Id’s of each Centre / Department
- The number of entries for each spreadsheet
- Any additional comments relating to the spreadsheets or to spreadsheets from Centres not supplied

13.2 CAL will acknowledge receipt of the spreadsheets from Monitoring Body and confirm or query any discrepancies in the number of spreadsheets provided, Centre / Department names / Id’s and the number of entries for each spreadsheet

13.3 The copying identity, that is the particular university and each Centre / Department are entered onto the CAL database. The following steps occur;

- The CAL licence number which identifies the university monitored is retrieved from the CAL database;
- The year of monitoring is added to the profile of the CAL licence number;
- Each Centre / department is allocated a CAL organisation number; and
- CAL organisation numbers are linked to the Monitoring Body Centre / department ID, the year in which the distribution relating to the records will occur, and the distribution pool (that is, “copied for” code).

14. Spreadsheet Data Preparation

14.1 The data will be loaded into the “staging/cleansing” area of CAL’s database where the data will be cleaned. . The cleaning includes;

- checking for empty fields
- deletion of characters not utilised by CAL
- field validity checks

14.2 Where there are anomalies CAL will clarify with Monitoring Body.

15. Lot Identification

15.1 Each spreadsheet will be sorted by University, Centre Code and Department Name (this ensures all data from the same area of the university ends up in the same batch in the database and that only one lot id is created for each department). The department names are entered into CAL's database ready for the spreadsheet to be imported.

16. Importing of Data

16.1 The Lot Id’s will be generated by the CAL database. The data from the spreadsheet pertaining to each Lot Id will be electronically imported into that Lot Id on the CAL database.

16.2 Electronic queries are made to ensure that each spreadsheet has been imported, and that the number of entries imported equals the number of spreadsheet entries.

17. Data Processing

17.1 The electronic Lots are assigned to Data Researchers. The process outlined in sections 18, 19, and 20 takes place.

18. Number of Copies/Target Audience

18.1 The “Number of Copies/Target Audience” data is entered in the database according to the following codes;

- For Distribution to Internal Students (I) = 002
- For Distribution to External Students (E) = 003
- For Distribution to Staff (S) = 007

18.2 There is no processing of bibliographic data for a single copy/communication made with the code “S”.

18.3 In the event that volumes in two or more of the fields “N,O, P, T,U and V” have been provided, the data researcher enters separate instances of copying/communication and the copies/communication will be treated in accordance with paragraphs 18.1 and 18.2.

18.4 If fields ‘N, O, P, T, U and V’ are empty, contain nonsensical entries or are ambiguous, the researcher refers to the surrounding records for possible clarification. For instance, a pattern of copying/communication for that person may already be established. When in doubt, the matter is referred to Monitoring Body or the university for resolution.

19. Volume

For all records with Type of Count W (other than Newspapers which are one page), change CAL Volume to Amount Copied or Communicated

by dividing by 300 and rounding up.

20. Material Type

20.1 The correct material type is allocated to each copying record in accordance with the definitions set out in section 9.

D. EXCLUSIONS

21. Copying and Communications to be excluded from processing

- 21.1 Copying and communication of material published by Universities Australia is excluded from processing altogether and does not form part of the data file sent to Monitoring Body.
- 21.2 Copying and communication of material published by the Department of Education Science and Training (DEST), including its predecessors DETYA, DEET and DEETYA is excluded from processing altogether and does not form part of the data file sent to Monitoring Body.
- 21.3 Copying and communication of material published by the High Court of Australia (including High Court decisions) is excluded, from the date notified, from processing altogether and does not form part of the data file sent to Monitoring Body.
- 21.4 Copying and communication of material for which CAL has received notification in writing from the Universities Australia that the Universities have been granted a copyright free licence is excluded, from the date notified, from processing altogether and does not form part of the data file sent to Monitoring Body.
- 21.5 Copying and communication of material for which CAL has received notification in accordance with clause 1.7.3 and Appendix 8 of the Monitoring System, and which contains sufficient details to allow accurate identification of copyright free works, is excluded from processing altogether and does not form part of the data file sent to Monitoring Body.
- 21.6 All entries marked as photocopying undertaken by students are excluded from processing altogether and do not form part of the data file sent to Monitoring Body.
- 21.7 All photocopying marked as ILL (Inter Library Loan) are excluded from processing altogether and do not form part of the data file sent to Monitoring Body.
- 21.8 All entries marked as photocopying for inclusion in examinations are excluded from processing altogether and do not form part of the data file sent to Monitoring Body.
- 21.9 When copies of the material copied have been provided by the university, any copying of quotations and extractions of three paragraphs or less are excluded from processing altogether and do not form part of the data file sent to Monitoring Body.
- 21.10 Works that have been protected by copyright, but for which that protection has expired, are known as works in the public domain. Works in the public

domain, are to be excluded from the data file sent to Monitoring Body. Guidance in determining whether a work is in the public domain is as follows:

Authors who died up to and including 31 December 1954:

Copyright expired on 31 December 2004.

Authors who died in 1955 (1 January-31 December) are still in copyright.

Copyright will expire at midnight on 31 December 2025 (=life +70 years)

No new material will fall into the public domain for 20 years

PHOTOGRAPHS

Photographs taken up to and including 31 December 1954:

Copyright expired on 31 December 2004

Photographs taken on or after 1 January 1955:

Copyright will last until 70 years after the death of the photographer

EXCEPTIONS

The copyright term for material in which the Government owns copyright remains 50 years

The copying of unpublished material under section 51(1) also remains.

If there is copying of works which are in the public domain based on the above principles, then the record is excluded from processing completely and does not form part of the data file sent to Monitoring Body.

If there is a new edition of a work (e.g. a modern translation or an annotated edition) or in the case of music, a new arrangement, then that work may be protected by copyright. Using all available knowledge and resources, the researcher should make a judgement as to the copyright protection of the work.

21.11 If it is obvious from the information contained in the record forms or spreadsheet that the copies for which the “copied for” code is S and for which the “material type” code is either L, LR or U, fall into the following category they are to be excluded from processing altogether and will not form part of the data file sent to Monitoring Body:

21.11.1 Section 182A of the Copyright Act exempts the copying of judgments and legislation from payment in the following circumstances:

- it applies to Federal or State legislation or subordinate legislation, judgments or awards of Federal or State Courts or Tribunals only, and does not include headnotes;
- for a particular person, for a particular purpose; and
- it must be a single copy.

E. DATA CHECKING AND AMENDMENT

22 Data Checking

22.1 CAL employs a data checker/auditor who is not a data researcher, to verify that the data researchers have processed in accordance with this protocol.

22.2 The auditor electronically retrieves each Lot I.D. and checks each entry line within that lot to ensure that:

- (i) No records have been duplicated or omitted. This is done by comparing the entry lines in CAL's database to the original records supplied by the university
- (ii) What has been entered in CAL's database is accurate. This is done by comparing columns on the record forms and fields of the spreadsheet to the corresponding fields in the database, as well as to any amendments made by data researchers. In respect of the hardcopy system the columns are
 - (a) Number of pages of original material
 - (b) Number of times each page copied
 - (c) Copied For
 - (d) Type
 - (e) Included Incidental artworks
 - (f) Publication Name
 - (g) Author
 - (h) Publisher
 - (i) ISSN/ISBN/ISMN
 - (j) Output Type
 - (k) Total Pages Copied

In respect of the Electronic Use System, the fields are;

- (a) Copied For
- (b) Material Type
- (c) Amount copied
- (d) Number of Copies
- (e) Coursepack

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

22.4 As a final check, system query reports are run at a macro level to ensure that all data is intact and that there are no invalid entities. In respect of any anomalies, the processes as set out in C are repeated.

23 Monitoring Body Audit

- 23.1 The data is audited by Monitoring Body. The scope and timing of the audit by Monitoring Body is set out in other documents.
- 23.2 Once the audit is completed and any corrections made, the electronic datafile is sent to Monitoring Body.

24 Other Monitoring Body Activity

- 24.1 Monitoring Body may, on the separate instruction and at the expense of CAL or Universities Australia as the case may be calculate and make a separate record of the volume of copies referred to in paragraphs 10.2 to 10.5, such record to be made available only to the instructing party.