

External Document

Catalogue Quality Control Procedures

UK Data Archive

Creator:	Data & Support Services/Information Development
Maintained by:	Data & Support Services/Information Development
Contributor:	
Version:	01.00
Controlled document No:	84
Last Amended:	2 March 2009
Review due date:	2 August 2009

Data & Support Services/
Information Development
UK Data Archive
University of Essex
Wivenhoe Park
Colchester
Essex
CO4 3SQ

email: help@esds.ac.uk
Tel: +44 (0) 1206 872001
Fax: +44 (0) 1206 872003
data-archive.ac.uk

Table of Contents

Document Control	3
1. Abbreviations	3
2. Definitions	3
3. Introduction	4
4. Quality control procedures prior to study release	4
4.1. Title check.....	4
4.2. Names authority list check.....	5
4.3. Subject categories	5
4.4. Access codes (pre-release)	5
4.5. Proof-reading and check for completeness.....	6
4.6. Checking of index terms	6
4.7. Error amendment	6
4.8. Recording pre-release catalogue record checks	6
5. Catalogue record release	7
5.1. Sending the release message	7
6. Post-release catalogue record quality control procedures	7
6.1. Outreach and Training section quality control check.....	7
6.1.1. Recording of the post-release check.....	8
6.1.2. Resolution of queries raised at the post-release check	8
6.2. Support Services check	8
6.3. ESDS Longitudinal check	8
6.4. ESDS Qualidata check	8
6.5. ESDS International Micro check	9
6.6. ESDS International Macro check	9
7. Updating catalogue records	9

Document Control

Version	Notes	Last Amended
external 01.00	First external version	2009-01-08

Replaces or supersedes:

Guide to Catalogue Quality Control Checks

Review terms:

Six-monthly

Is related to:

Cataloguing Procedures and Guidelines

Cataloguing Tips

Data Processing Procedures

Dataset release checklist

Indexing Procedures (in development)

Policy Procedures and Guidelines for Publicity and Web Sites

Recording Changes to Catalogue Records

Note to external readers

This document covers the catalogue quality control procedures in use at the UK Data Archive (UKDA). For external readers, it is for information only, and readers should note that some of the documents referenced above may not be available for external use.

1. Abbreviations

After the first instance in this document, sections/teams are referred to by abbreviations in common use at the UKDA. However, job titles are always provided in full for clarity.

2. Definitions

Arc-release

The 'arc-release' list is an internal subscription email list used to alert UKDA and ESDS staff and others to new datasets and new editions released. It is controlled by the University of Essex Computing Service.

CALM database

The CALM database is the internal UKDA database used to track study ingest processing progress.

Esds-all

The 'esds-all' list is the Jiscmail list used to alert external ESDS users and others to new datasets and new editions released.

Humanities and Social Science Electronic Thesaurus (HASSET)

HASSET is a subject thesaurus which has been developed by the UKDA over the past 20 years. Initially based on the 1977 edition of the UNESCO thesaurus, HASSET has been continuously expanded and updated for use in its online retrieval system.

Processor

The 'Processor' is the member of UKDA staff responsible for all or part of the ingest processing of a study, including cataloguing and indexing. The Processor will usually be a member of the Data Services (DS), Qualidata, Census or History Data Service (HDS) section. Where updates to existing catalogue records are described in section 5, the Processor is the member of staff who undertakes the catalogue update work, irrespective of whether they are the original ingest Processor of said study.

Service Level Definition (SLD)

The SLD for ingest processing is agreed between the ESDS and its funders and governs the period in which a study must be processed and released for secondary use, once its acquisition is deemed 'complete' (i.e. all items of data, documentation and metadata have been received). For quantitative studies, the current SLD is 30 working days. For qualitative studies, the SLD is 35 working days. Enhanced quantitative processing, i.e. Nesstar addition and release, is subject to an additional SLD of 35 working days after dataset release.

3. Introduction

This document governs quality control procedures for all UKDA records included in the Data Catalogue. For all studies, certain catalogue record checks are required at various stages of processing before the study can be finally released to the Data Catalogue.

4. Quality control procedures prior to study release**4.1. Title check**

When a new study is passed to the DS team for ingest processing, the member of DS staff who completes the allocation of a new study number and title (see the *Data Processing Procedures* document) will notify the Metadata and Resource Discovery Officer via email of the existence of the new study and, if appropriate, the location of its red folder and associated files. The Metadata and Resource Discovery Officer will then check that:

1. the study title (and alternative title where appropriate) adheres to the rules laid down in the *Cataloguing Procedures and Guidelines* document;
2. the alternative title (where included) is correct and appropriate;
3. there is consistency of title across generic groups and series;
4. the existence of possible related studies within the Data Catalogue.

The title checks will normally take place prior to (or in the early stages of) ingest processing, so that the correct title is available in the CALM database Read and Note file sections. It should be noted however that further information about a particular study to supplement the catalogue record may only become apparent as ingest processing progresses. Therefore, checks 2, 3 and 4 may be made at the later proof-reading check stage instead. If additional

information is discovered during processing that may mean the addition of an alternative title, or link the study to an existing generic group or series, the Processor should inform the Metadata and Resource Discovery Officer of this when the proof-reading request is made (see section 3.1.5 below) so that informed checks may be performed.

New editions

Where the study is a new edition, and changes to the title are complex, the Processor should seek advice from the Metadata and Resource Discovery Officer as to whether a check is required.

Recording the title checks

The CALM database will eventually have a checkbox to record title checks, but this facility is not yet available.

4.2. Names authority list check

If sections of the catalogue covering Depositor, Principal Investigator, Sponsor (or any of the other 'name' fields) require the addition of a new name/organisation to the Names Authority list, this should be added according to the rules laid down in the *Cataloguing Procedures and Guidelines*. The Metadata and Resource Discovery Officer will receive automatic notification to 'approve' the new addition. In case of doubt, the Processor should seek advice from the Metadata and Resource Discovery Officer **before** addition of the name. The new version of the catalogue input programs currently in development will also provide opportunities for consistency checks across series.

4.3. Subject categories

The Metadata and Resource Discovery Officer will check the subject categories allotted to the study to ensure that the main subject(s) of the study are covered, and that major studies include the appropriate subject category. This check will be performed at the proof-reading stage once the catalogue record is complete (see section 3.1.5 below). If the available subject categories do not cover the new study adequately, the Processor may request the addition of a new subject category via discussion with the Metadata and Resource Discovery Officer, who will in turn consult the Metadata and Resource Discovery Manager. (Addition of a new subject category to the existing list cannot always be done instantly and should not hold up initial release of a study if users are waiting or there is time-sensitivity with regard to the SLD; the new category may be added post-release.)

4.4. Access codes (pre-release)

If the study is part of a series, the Metadata and Resource Discovery Officer will check the access code allotted to the study to ensure that there is consistency over the series. This can be done at the proof-reading check stage (see section 3.1.5 below).

If the series is a long-standing one, access codes may have changed over time according to amendments made to the UKDA/ESDS catalogue systems, for example a code change from '2' to '2003A' is a system update, and denotes the same access conditions. Whilst the UKDA Acquisitions section have done considerable work with depositors to ensure that updated access conditions may be applied to a whole series, queries will sometimes arise. Such discrepancies should be reported to the Processor and the Data Services Manager, who will notify the Acquisitions section to check with the depositor. If the access condition is accurate according to the new study licence form and seems logical over a series (for example the change from '2' to 2003A) this should not hold up initial release of the study. Subsequent post-release amendments either to the new study, or the whole series, may be made as necessary.

There may also be rare occasions when certain studies in one series may be subject to differential or conditional access, or one wave is unavailable. Again, these discrepancies should be reported to the Data Services Manager, in case circumstances have changed over time.

4.5. Proof-reading and check for completeness

The catalogue record must be checked for errors by a senior member of the DS team (the DS proof-reader) before the study is released. The DS team may decide between themselves, subject to approval by the Data Services Manager, who will check each others' work; this system allows for flexibility according to staff working hours and during holiday/sickness periods.

To start the proof-reading and check, the Processor will notify **both** the appropriate DS proof-reader and the Metadata and Resource Discovery Officer via email. At the same time as DS proof-reading, the Metadata and Resource Discovery Officer will undertake checks on the subject category, access code, and any elements of the title check that could not be completed earlier (see section 3.1.1). (In the absence of the Metadata and Resource Discovery Officer, the Data Services Manager or Metadata and Resource Discovery Manager should be consulted.)

The DS proof-reader must read the complete catalogue record, including generic information as appropriate, with reference to the rules laid down in the *Cataloguing Procedures and Guidelines*, *Cataloguing Tips* and *Dataset release checklist* documents where necessary. Index terms should also be checked in consultation with the Processor (see section 3.1.6). In case of query, the Data Services Manager or Metadata and Resource Discovery staff may be consulted for advice.

For Census and HDS studies, the Census or HDS staff will perform the pre-release proof-reading check within their sections.

4.6. Checking of index terms

Once the catalogue record has been assessed, the index terms should be checked as part of the pre-release proof-reading (see 3.1.5 above) with reference to the *Dataset release checklist*. The Processor should also be reminded to check that index terms provide adequate coverage of the study. Keyword indexes must be compiled in accordance with the *Indexing Procedures* document (when available) and should include a date and geographical location. In case of query, the Data Services Manager may be consulted. If concepts included in the study are not adequately covered by the current HASSET thesaurus, suggestions for new index terms by Processors are welcome. These must be discussed with the Metadata and Resource Discovery Officer.

4.7. Error amendment

When the proof-reading checks are complete, the DS proof-reader and the Metadata and Resource Discovery Officer will notify the Processor via email. Any edits deemed necessary must be made to the mutual satisfaction of all three parties before the study is released. The proof-reading checks and edits must be completed in timely fashion to ensure that the processing SLD requirement is met.

4.8. Recording pre-release catalogue record checks

The DS proof-reader should record the date catalogue checks are completed, and the time taken to complete checks ('Catalogue Checks I') on the appropriate pages in the CALM processing database. Completion of these sections will move the CALM entry from the 'Active Log List' to the 'Completed Log List' in the database, which is important for the recording of the SLD. In future, CALM will have more reporting facilities for catalogue checks, but these are not yet available.

5. Catalogue record release

Once the catalogue record and index have been checked and any reported errors amended, and the Processor has ensured that the study is correctly archived on the UKDA preservation server, the catalogue record may be released to the main UKDA/ESDS catalogue. This may be done by the Processor if able, but otherwise may be done by the Data Services Manager or a senior member of the DS team.

5.1. Sending the release message

Once the catalogue record is released, the Processor must send a release message to the arc-release email subscription list, to notify internal staff that a new dataset/edition has become available. Processors should however be aware that the catalogue record documentation table and the web catalogue index update overnight following the study's release. Therefore, if the study is particularly complex, or is a new edition, it may be considered prudent to delay this message until the day after release to ensure that update has completed correctly. The catalogue record will still be available from the moment of release on the New Data Released web pages and, in the case of a major study, in the Major Studies table.

Once the catalogue record has been released, it becomes subject to post-release catalogue quality control procedures as appropriate, which must take place within 14 days of release.

6. Post-release catalogue record quality control procedures

All catalogue records receive a pre-release quality control check as detailed above. The checks detailed below are extra, and are not a substitute for the pre-release check. They will only be performed on those studies that are chosen by random selection, or meet the criteria outlined.

6.1. Outreach and Training section quality control check

The Outreach and Promotion Officer, based in the UKDA Outreach and Training (O & T) section, will perform a weekly catalogue record quality control check. This check comprises the selection of **one** newly-released catalogue record for proof-reading from the weekly 'New releases and new editions' email. This email is circulated every Monday to the ESDS-all Jiscmail list by the O & T section, based on information taken from the 'New data releases in the last 7 days' web page on the ESDS web site (allowing for a cut-off date of the previous Thursday).

Catalogue records for the post-release quality control check must be selected in rotation to cover the ESDS service (Core, Government, Longitudinal, Qualidata and International), Census and HDS studies. If no studies have been released that week for the appropriate rotational choice, one catalogue record should be chosen at random from the available releases. This quality control check should be completed within 14 days of the study release, and comprises:

- proof-reading of the selected catalogue record with reference to the *Cataloguing Procedures and Guidelines* document. Reference may also be made to the *Policy Procedures and Guidelines for Publicity and Web Sites*, though it must be borne in mind that appropriate formats for web pages are not always correct for catalogue records, which adhere to international archival standards in the first instance;
- a check to ensure that appropriate index terms are included to cover date and geographical location.

Any errors found must be marked and reported to the Processor responsible for compilation of the catalogue record. The Processor will discuss any changes to be made with the Outreach and Promotion Officer, and complete them as appropriate.

The Outreach and Promotion Officer should advise the Data Services Manager of the catalogue record selected for checking each week. This will help to streamline and co-ordinate the

checking process where the selected record is subject to service checks, such as those done for ESDS International or ESDS Qualidata studies (see sections 6.3 to 6.5 below).

6.1.1. Recording of the post-release check

Once any errors have been fixed to the mutual satisfaction of the Processor and the Outreach and Promotion Officer, the Outreach and Promotion Officer will record the time taken to check in the 'Catalogue Checks II' section of the 'ABC' page of the CALM database entry for the study concerned. In case of query, the Outreach and Promotion Officer will consult the Data Services Manager.

6.1.2. Resolution of queries raised at the post-release check

The deciding criteria in case of query is whether the Processor feels that the information provided will be logical and clear to data users. If the query needs further discussion, the Processor and Outreach and Promotion Officer will consult the Data Services Manager, who may in turn consult the Metadata and Resource Discovery Manager, whose decision is final.

6.2. Support Services check

On occasions, studies acquired for the UKDA/ESDS collection may be subject to special or unusual access conditions. These may include Special Licence Access studies, or those with conditional or differential access for parts of the study. In these cases, the DS section will advise the UKDA Support Services (SS) section of the study concerned and its unusual access arrangements prior to release, so that discussions on the content of relevant sections of the catalogue record may take place as necessary. Once processing is complete, the DS section will then advise the SS section of the study's subsequent release via the arc-release email list. The SS section will then check the wording of the catalogue record to ensure that the access conditions are meaningful and workable within the UKDA user support and dataset supply systems. This check should be performed within 14 days of the study release, and any change requests must be reported back by email to the DS section, who will make the changes and re-release the catalogue record. Eventually, there will be a facility to record this check in the CALM database.

6.3. ESDS Longitudinal check

The Service Manager, ESDS Longitudinal (or their delegate as advised to the DS section), will perform a post-release check on all new releases and new editions of studies categorised as part of the ESDS Longitudinal collection. These checks should be performed within 14 days of the study release as advised by the Processor's release message, and will include (but not be limited to) proof-reading the catalogue record to ensure that the methodology and other characteristics of the study are correctly described within the catalogue record, and that the time dimensions of the study are accurately noted. Any errors should be reported back to the Processor, amendments made and the catalogue record re-released. In cases of query, the Data Services Manager should be consulted.

The free-text section of the 'Processing' screen of the study's record in the CALM database may be used to record completion of the ESDS Longitudinal post-release check.

6.4. ESDS Qualidata check

The Senior Qualitative Data and Support Services Officer (or their delegate as advised to the DS section), will perform a post-release check on all new releases and new editions of studies categorised as part of the ESDS Qualidata collection. These checks should be performed within 14 days of the study release as advised by the Processor's release message, and will include (but not be limited to) proof-reading the catalogue record to ensure that the methodology and

other characteristics of the study are correctly described within the catalogue record. Further checks include ensuring that all the documentation is available and that all hyperlinks work. Any errors should be reported back to the Processor, amendments made and the catalogue record re-released. In cases of query, the Data Services Manager should be consulted.

The free-text section of the 'Processing' screen of the study's record in the CALM database may be used to record completion of the ESDS Qualidata post-release check.

6.5. ESDS International Micro check

The Service Manager, ESDS International (or their delegate as advised to the DS section), will perform a post-release check on all new releases and new editions of studies categorised as part of the ESDS International collection. These checks should be performed within 14 days of the study release as advised by the Processor's release message, and will include (but not be limited to) proof-reading the catalogue record to ensure that the methodology and other characteristics of the study are correctly described within the catalogue record, and that the cross-national geographical coverage of the study is accurately noted. Any errors should be reported back to the Processor, amendments made and the catalogue record re-released. In cases of query, the Data Services Manager should be consulted.

The free-text section of the 'Processing' screen of the study's record in the CALM database may be used to record completion of the ESDS International Micro post-release check.

6.6. ESDS International Macro check

The Outreach and Promotion Officer will proofread all new and significantly updated International Macro catalogue records within 14 days after their release (they will receive the pre-release checks as normal). When notification of a new catalogue record is received, or of significant updates or additional index terms required for an existing International Macro record, the Processor will email the Outreach and Promotion Officer, who will subsequently proofread the new record/updated information and additional index terms. This constitutes a different kind of check, performed against the original information provided by ESDS International for new records and significant updates. These materials (often a Word document and/or Excel file) must be archived with the study under SN/noissue/original/.

7. Updating catalogue records

Whilst new editions of studies will be subject to the checking procedures described above, occasionally catalogue records may need to be updated. If the changes required are very slight, and the Processor is confident that their amendment is correct, the catalogue record may be re-released. However, if changes required are complex, or involve more than one section of the catalogue record, the Data Services Manager, or a senior member of the DS team, should be consulted regarding wording and before re-release.

Significant changes, such as the amendment of access codes, or those following email correspondence with the study depositor, should be recorded in the study Note file (a new version will need to be plattered), and the email correspondence added to the study red folder. In these cases, the procedures laid out in the *Recording Changes to Catalogue Records* document should be followed.