



The work is licensed under the Creative Commons Attribution-Non-Commercial-Share Alike 2.0 UK: England and Wales Licence. To view a copy of this licence, visit creativecommons.org/licenses/by-nc-sa/2.0/uk/

Catalogue Quality Control Procedures

Version 2.0

PUBLIC VERSION

12 October 2010

T +44 (0)1206 872001

E help@esds.ac.uk

www.data-archive.ac.uk



UK DATA ARCHIVE
UNIVERSITY OF ESSEX
WIVENHOE PARK
COLCHESTER
ESSEX, CO4 3SQ

WE ARE SUPPORTED BY THE UNIVERSITY OF ESSEX, THE ECONOMIC AND SOCIAL RESEARCH COUNCIL, AND THE JOINT INFORMATION SYSTEMS COMMITTEE

Contents

1.	Resource Discovery Section (RDS) quality control checks (pre-release)	3
1.1.	Title check	3
1.2.	Names authority list check	4
1.3.	Subject categories.....	4
1.4.	Access codes (pre-release)	4
1.5.	Proofreading and check for completeness.....	4
1.6.	Checking of index terms.....	5
1.7.	Error amendment	5
1.8.	Recording pre-release catalogue record checks.....	5
2.	Catalogue record release	5
2.1.	Sending a release email message	5
3.	Post-release catalogue record quality control procedures	5
3.1.	RDS post-release checks.....	5
3.2.	Outreach and Training section quality control check	6
3.2.1.	Recording of the post-release check	6
3.2.2.	Resolution of queries raised at the post-release check.....	6
3.3.	Studies with special conditions – notifying Support Services	6
3.4.	ESDS Longitudinal check.....	7
3.5.	ESDS Qualidata check.....	7
3.6.	ESDS International Micro check	7
3.7.	ESDS International Macro check	7
4.	Updating catalogue records	8

Scope

This document details quality control procedures for all catalogue records included in the UK Data Archive/Economic and Social Data Service (ESDS) 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.

It should be noted that some of the documents referenced within the text below are not publicly available, but external readers may of course contact the Archive in case of query.

Definition of Terms

Arc-release

The 'arc-release' list is an internal subscription email list used to alert Archive 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 Archive database used to track study 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 Archive 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 Archive staff responsible for all or part of the ingest processing of a study/data collection, including cataloguing and indexing. 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 the study /data collection.

Resource Discovery Services (RDS) HelpDesk

The RDS HelpDesk is an internal Archive system for queries to RDS, administrated via the JIRA HelpDesk system.

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 data collections, 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.

Storm database

The 'Storm' database is the internal Archive database used to track study processing progress.

1. Resource Discovery Section (RDS) quality control checks (pre-release)

1.1. 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 notifies RDS via the JIRA-based Resource Discovery HelpDesk of the existence of the new study. The RDS will then check:

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.

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 proofreading 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 RDS of this when the proofreading request is made (see below) so that further 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 RDS as to whether a check is required.

Recording the title checks

Completed title checks are recorded via the RDS HelpDesk.

1.2. 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 in the catalogue input program according to the rules laid down in the *Cataloguing Procedures and Guidelines*. The RDS will receive automatic notification to 'approve' the new addition. In case of doubt, the processor should seek advice from RDS before addition of the name.

1.3. Subject categories

The RDS 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 proofreading stage once the catalogue record is complete (see 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 RDS. (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.)

1.4. Access codes (pre-release)

If the study is part of a series, RDS will check the access code allocated to the study to ensure that there is consistency over the series. This can be done at the proofreading check stage (see below).

If the series is a long-standing one, access codes may have changed over time according to amendments made to the Archive/ESDS catalogue systems; e.g. a code change from '2' to '2003A' is a system update, and denotes the same access conditions. Whilst the Archive's Acquisitions section have done a 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 (e.g. 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.

1.5. Proofreading and check for completeness

When complete, processors should ask another member of the DS team (the DS proofreader) to check the catalogue record 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 proofreading and check, the Processor will notify the appropriate DS proofreader by email and the Resource Discovery Section via the RDS HelpDesk. At the same time as DS proofreading, RDS will undertake checks on the subject category, access code, and any elements of the title check that could not be completed earlier.

The DS proofreader should read the complete catalogue record, including generic information as appropriate with reference to the rules laid down in the *Cataloguing Procedures and Guidelines* document where necessary. Index terms should also be checked in consultation with the processor. In case of query, the Data Services Manager or RDS section may be consulted for advice.

For Census and HDS studies, Census or HDS staff will perform the pre-release proofreading check within

their sections.

1.6. Checking of index terms

Once the catalogue record has been assessed, the index terms should be checked as part of the pre-release proofreading. Index terms should provide adequate coverage of the study. Keyword indexes must include a date index term 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 RDS and then requested via the RDS HelpDesk.

1.7. Error amendment

When the proofreading checks are complete, the DS proofreader will notify the Processor via email and RDS via the RDS HelpDesk. Any edits deemed necessary should be made to the mutual satisfaction of all parties before the study is released. The proofreading checks and edits must be completed in timely fashion to ensure that the processing SLD requirement is met.

1.8. Recording pre-release catalogue record checks

The DS proofreader should record 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.

2. 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 Archive preservation server, the catalogue record may be released to the main Archive/ESDS catalogue. This may be done by the processor if able, or the Data Services Manager/senior member of the DS team.

2.1. Sending a release email message

Once the catalogue record is released, the processor should send a release message to the 'archive-release' email subscription list, to notify internal staff that a new study/new 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 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 should take place within 14 days of release.

3. 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.

3.1. RDS post-release checks

The RDS section perform post-release checks on all newly released studies, following notification via the archive-release list. Post-release checks are performed the day following release, or as soon as possible afterwards. Checks include:

verification that the catalogue record can be retrieved online via the Data Catalogue.

Where studies are part of a major series, checks also include verification whether the study is included in the major studies table on the appropriate major studies web page.

Where the release is a new edition, a check is to ensure the new edition statement is included in the catalogue record and Read file (the Read file may not update until the day following release, due to the overnight update cycle of internal Archive systems).

The RDS post-release checks are logged in the RDS HelpDesk. If further action is required, the processor who released the study is contacted, or, in their absence, the Data Service Manager is notified.

3.2. Outreach and Training section quality control check

The Outreach and Promotion Officer, based in the Archive's 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 below).

3.2.1. 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.

3.2.2. 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, whose decision is final.

3.3. Studies with special conditions – notifying Support Services

On occasions, studies acquired for the Archive/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 Archive's 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, and recording of special access conditions may be done in the Storm database. As an extra check, the processor may ask the SS section to check the wording of the catalogue record after release, to ensure that the access conditions are meaningful and workable within the UKDA user support systems.

3.4. 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.

3.5. 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 qualitative data collections. 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.

3.6. 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.

3.7. 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 should 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 macro records and significant updates. These materials (often a Word document and/or Excel file) should be archived with the study under SN/noissue/original/.

4. 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.