

October 2021 DCS Reviews: Summary Report

Graham Parton (CEDA)

18th January 2022

Summary

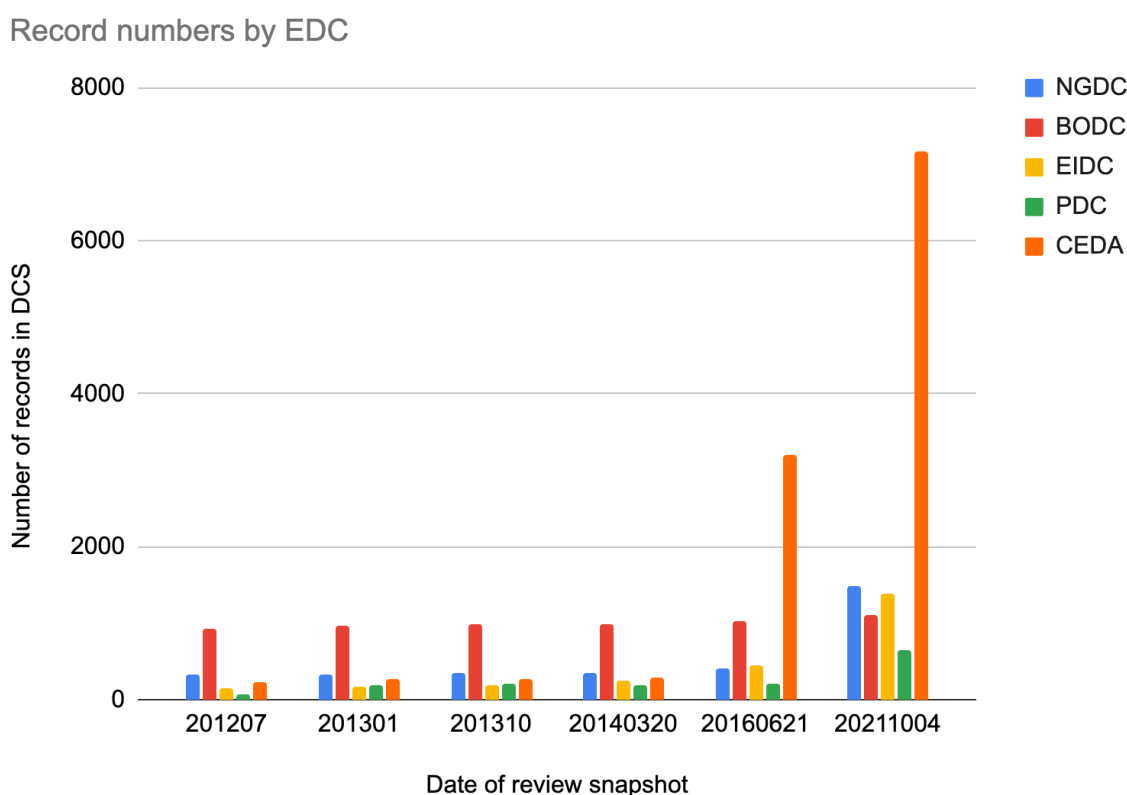
Comparison of the results of the latest round of metadata checks with previous rounds has revealed that despite a doubling of metadata records for the EDS as a whole since the last review round in 2016, metadata quality has been maintained at a high standard.

Having adapted the review process slightly from the 2016 round it was possible to isolate data centre records coming from collections of similarly templated records. Consequently, this review was able to give further insight as to the breakdown of record quality by record type. From this the strength of taking templating approaches to metadata record creation was demonstrated with resulting records being of higher quality overall, though there are some variance in results still. Additionally, non-geographic datasets were examined for the first time as an identifiable class of records within the sample, which showed a slightly lower than average scores - but again with a large spread of metadata quality scores.

Methodology

A snapshot of the DCS catalogue was taken 4th October 2021. At the time there were over 11,000 metadata records in the catalogue. This is a substantial increase in holdings since the previous DCS round of checks in June 2016 when there were 5304 records. Over this time period all data centres except the BODC have seen a doubling in the number of records produced.

The breakdown of records by issuing centre is given below for each round of DCS reviews, including the latest is given below:



In previous rounds the breakdown of record numbers by “data centre” had indicated various non-NERC data centre institutes being listed in conjunction with records within the NDC (e.g. Plymouth Marine Laboratory, Isle of Man Government Laboratory, University of Glasgow, Department of Geographical and Earth Sciences, Nirex and ARUP - full list in appendix A of the 2016 report). However, as with the reviews in previous rounds the focus was on the metadata record issuing centre which corresponds to the main NERC data centres.

The procedure for the reviews was adapted slightly from the June 2016 methodology to address a number of issues raised by the reviewers following the 2016 round of reviews. The main changes from the 2016 methodology were:

- Slight alteration to the review questions to make them easier to handle (e.g. turning the occasional negatively worded question into a positive one to match the majority of review questions)
- Splitting out the section on limitations into 2 separate sections to reflect GEMINI2.3: one for legal constraints (i.e. around end-usage licensing) and the other for access constraints (i.e. limitations on accessing the resource or not). Scoring has been adjusting to take this into account.

- c) Additional response options of 'not applicable' and 'unsure' were introduced to questions to further assist the review process. Scoring and reporting were adjusted accordingly to reflect their inclusion.
- d) Identifying collections of similarly structured records for each data centre from which only a handful of sample records would be taken for review allocations. This was to address cases where large numbers of similar records (examples such as CEDA's > 2000 records from model intercomparison projects; and the majority of BODC cruise dataset records). These will be referred to as 'collection' dataset records or 'templated' records in this report.
- e) No attempt was made to force the seeding of the top of the review lists to include a sample of records covered by previous review rounds as this had been identified to lead to an over-sampling of BODC records being reviewed in the 2016 review.
- f) Reviewer fatigue - the 2016 review round took place over the course of 1 day and reviewers flagged that they were making increasing mistakes towards the end. To address this the reviews were permitted over a 3 week period for reviewers to dovetail between their day-to-day work with a suggested target of ½ to 1 day's effort dedicated to the reviews.

As with the previous round of reviews all records within scope of the review for the EDS data centres were taken as available for review. This included datasets and series level records as in previous reviews but also 'non-geographic' datasets. This latter category is specifically identified within the NERC Data Catalogue Service and does account for a sufficient number of records to be included in the review. Record types not in scope included those marked as: 'Service', 'service-view', 'application', 'Dataset Collection*', 'service-download' and 'Model' - less than 100 records between these categories. (* these are not the same as the 'collection' dataset records covered in 'd' above, but akin to 'series' level records but not included as they related to ADS records which were not within scope).

Lists covering available records (or samples of collections of records as noted in d above) were then allocated to be reviewed amongst the review team, ordered to ensure that:

- i) a reviewer would not review any records from their own data centre.
- ii) The initial set of each list would include records from the samples of 'templated' records to be reviewed.

This latter point was essential to ensure that reviews of these sample records were completed at least, with the intention of then taking these to be representative of other records from their collections. Thus, as will be presented later, the scores for these sample records could then be applied to all members giving, in effect, a greater indication of the overall quality of the records. Similarly, averages of 'non-collection' records may also be indicative of other records from a given data centre. Note, however, that such scaled scorings must be taken with a degree of caution and can

provide extra challenges when attempting to characterise any progress (or not) from previous review rounds.

Review procedure

Reviewers were asked to view the records and complete a Google Form covering the same series of simple yes/no style questions and free-text feedback boxes to establish the value quality of the metadata and relay findings back to the record's originating data centre as used in earlier review rounds. Two additional options of 'unsure' and 'not applicable' were also provided to aid the review process. These were added following feedback from those involved in the 2016 review round. Please see later notes on how such responses were handled with review scoring.

This round utilised the full set of DCS review questions as used in review rounds 1 to 3 and the last round (round 5) in 2016 - round 4 only focused on 6 key questions.

The full set of checks each record was reviewed against are given in appendix B.

Using the checklist, errors in the metadata were ranked as for round 3:

Major there are problems affecting comprehensibility/usability of the metadata and the record needs attention as soon as is practicable

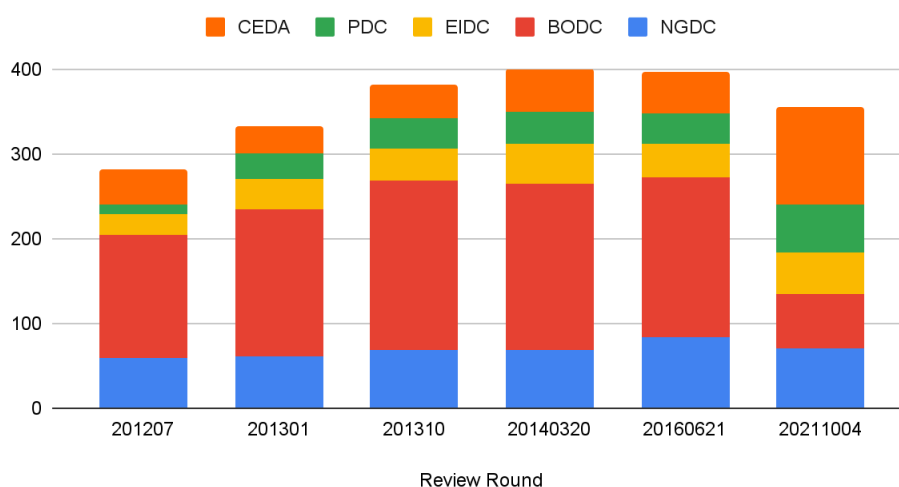
Minor field is considered to be lower priority for users discovery, usability and understanding of the resource

Reviewed Records

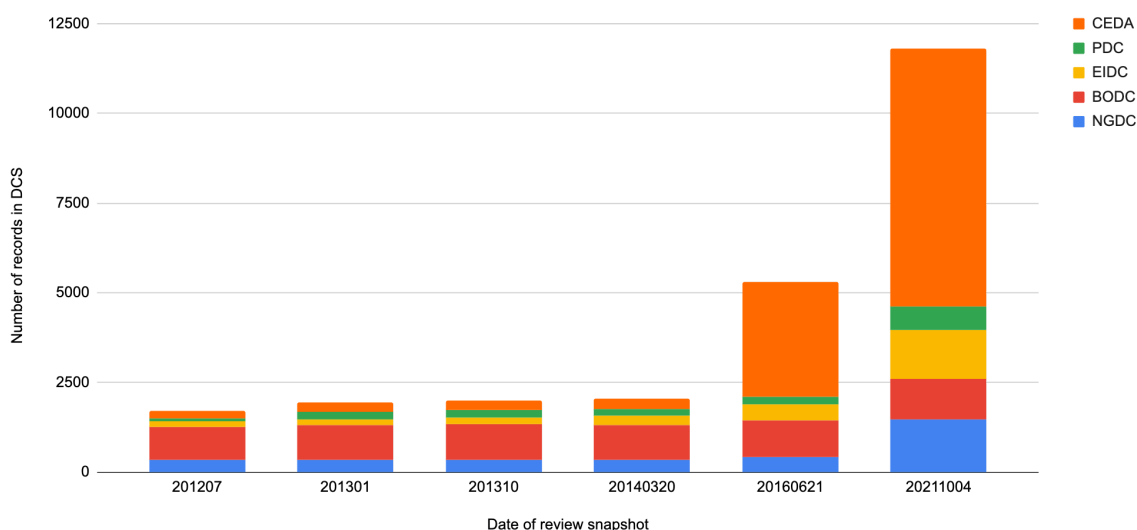
358 reviews were submitted during the review period, returning a sample size of just over 3% of all records within scope of the review. These numbers are lower than the previous review of 409 records which returned a sample size of around 7.5% of the DCS records in June 2016.

Overall, the number of records reviewed in the latest round of reviews lowest since the 2nd review round. However, the limited scope of the 4th round should be kept in mind and a more accurate comparison would be with round 3 and 5. Still, whilst the previous report indicated that a review day may be a good way to target this review procedure, allowing reviewers to spread their work over an extended period of time unfortunately led to a decrease in overall numbers of records being reviewed. Given 2016 reviewers' feedback about the draw-backs of such an intense review day and that feedback was generally positive to allow the reviews to be spread over a number of weeks, this may be a reduction that has to be accepted.

Data centre records reviewed per review round



Data Catalogue Service data centre record numbers for each review round



Given the large variation in numbers of records by data centre the 3% sample rate is indicative of the overall DCS records only, being much lower for CEDA records in particular given their large numbers as shown in table below:

Review Round	201207	201301	201310	20140320	20160621	20211004
NGDC	17.5	18.1	19.2	20.1	19.9	4.7
BODC	15.5	17.8	20.4	19.8	18.4	5.8
EIDC	16.0	20.6	19.9	19.3	9.0	3.5
PDC	16.9	14.7	17.6	19.9	16.4	8.9
CEDA	18.6	12.0	14.4	17.2	1.5	1.6

Sample percentages per round for each data centre.

This table demonstrates that there has been a decreasing sample size for each data centre in recent review rounds, which is unavoidable given the manual nature of the process and the rapid growth of record numbers. **This is an important caveat that needs to be considered in how robust these figures will remain and also how the review process will need to continue to adapt to accommodate this fact, especially when reporting changes in quality over time.**

The variability of sample representativeness by data centre can be explained in part the following factors:

- 1) 4 out of 13 reviewers were from CEDA, thus there was a greater available effort to review records from non-CEDA data centres
- 2) Effort per data centre/reviewer was not evenly distributed
- 3) Record list creation was done on a round robin basis, thus close-to-equal seeding of the review lists provided to each review would initially be relatively equally spread amongst the data centres for which they were reviewing. I.e. the allocation was not weighted to reflect the differing record volumes from each data centre.

Record Types

Within the DCS there are two levels of records permitted: Series and Datasets. Previous DCS reviews had assumed that all records were Datasets, whilst in actuality all CEDA's records had been issued at the Series level. Additionally, 'dataset' level records can be further split into 'datasets' and 'non-geographic' datasets.

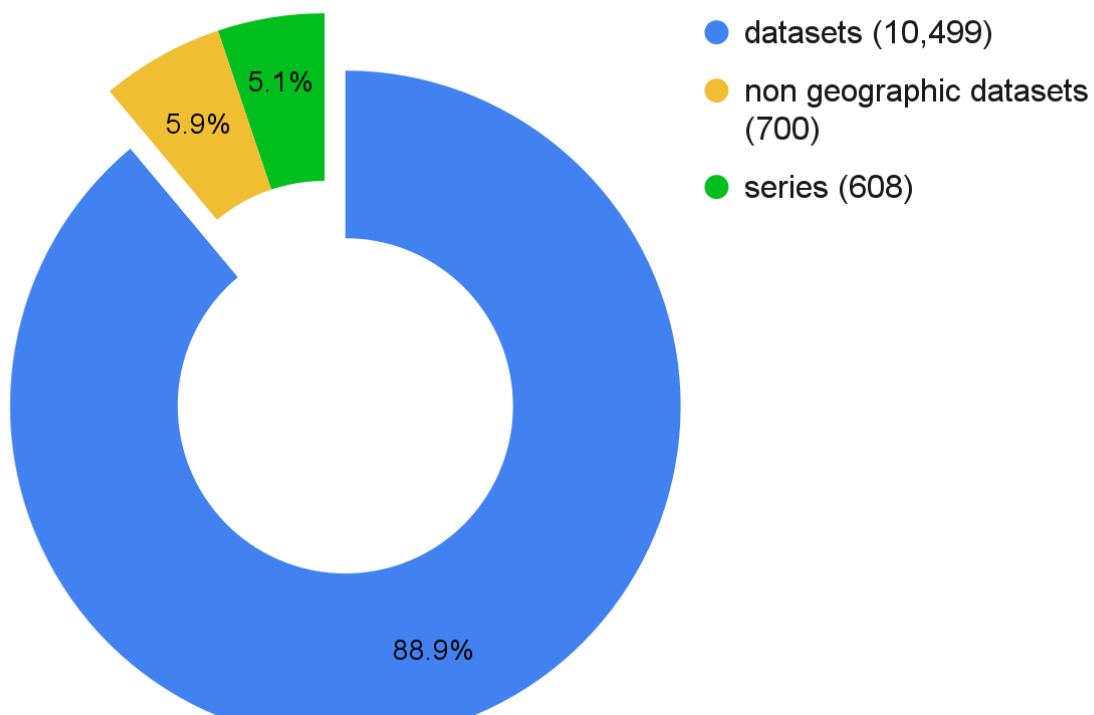
Given the nature of Series level records being a broader grouping of sometimes diverse component datasets, it was noted previously that the DCS metadata guidelines and reviews should take this into account. It had been suggested that guidelines for Series level records should have a freer scope than those for Datasets. However, no specific guidelines have since been drawn up with regards to the differences between these two record levels and thus the same review criteria have been applied to both levels of record. Note, however, that since the 2016 round of reviews CEDA have made specific efforts to address handling of series level records within the GEMINI2.3 standard. Within GEMINI2.3 the fields within the series and dataset records are very similar in scope with the nature of the series level content being less specific than dataset level records.

Additionally, within the reviews where records came from large collections of similar records, as mentioned earlier, the selected 'samples' were tracked through the review process - these will be referred to as 'sample datasets' below.

Consequently, it is possible to distinguish the differences in record quality by the following categories:

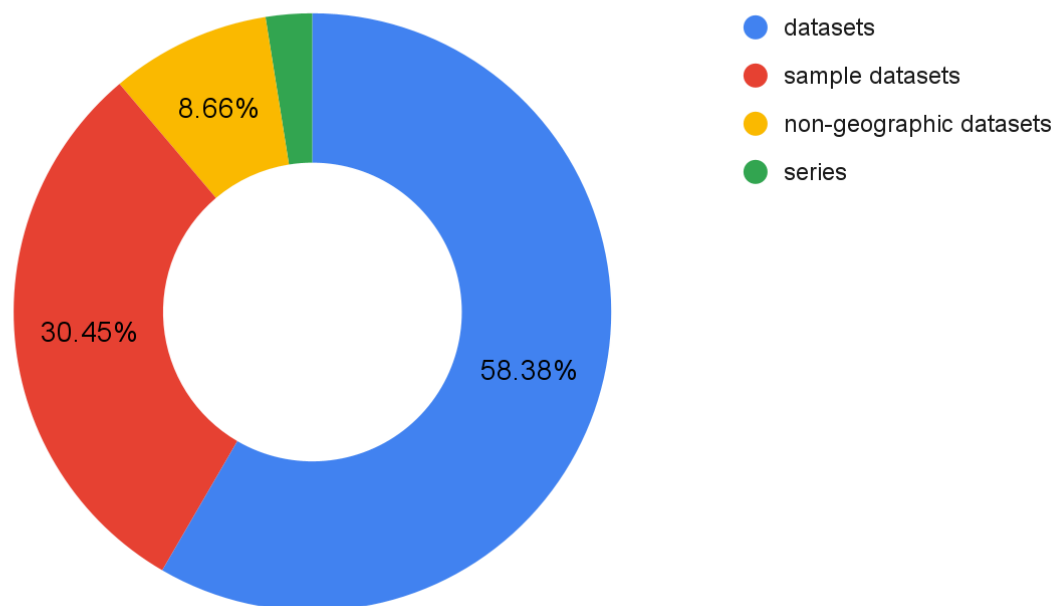
- Record type: dataset, non-geographic dataset, series
- 'Sample' datasets to represent where record templating is in use for record construction

However, it should be noted that for the entire DCS record population only 5.1% of records (608) are Series level and these issued only by CEDA (~11% of all CEDA's records in the DCS). Non-geographic datasets account for close to 6%, with EIDC being the main provider of such records.



Within the sample the split by record type was as follows:

Sample percentages by record type



Whilst the datasets/sample datasets are in proportion to the dataset records of the data catalogue service, there was a higher proportion of non-geographic datasets within the sample and fewer series records. The low numbers of series records, just 9 of the 358 reviews compared to nearly 800 in the catalogue service, means that any associated review content should only be applied to the records reviewed and much greater care taken when taking these results as representative to other series level records..

Results

Checks conducted on the comprehensibility of metadata are by their very nature subjective but ensuring checks are conducted by non-specialists helps to fulfil the brief that metadata should be understandable by a first year undergraduate.

Technical issues

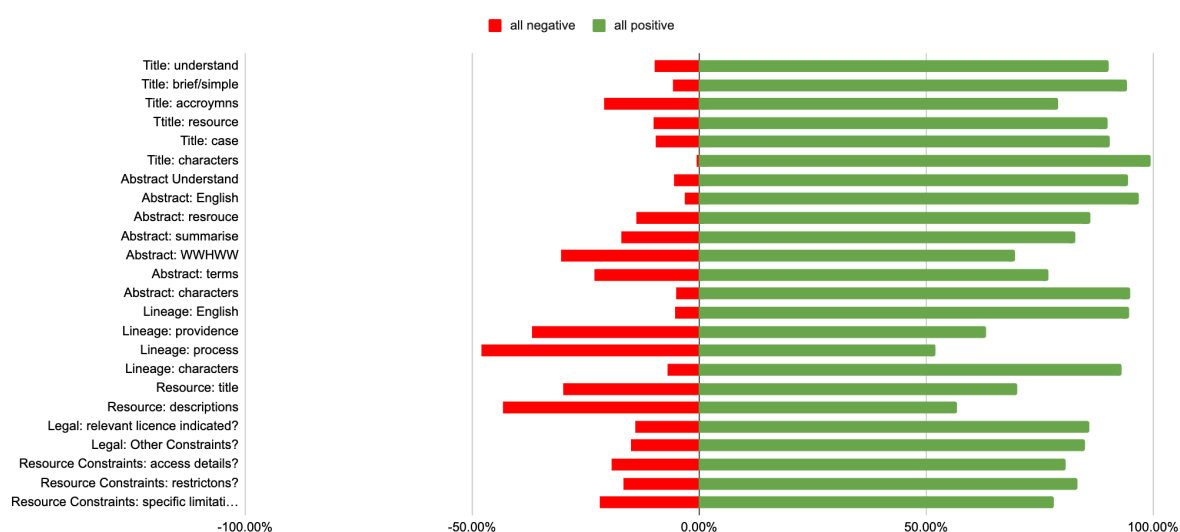
During the review process a number of technical issues were either reported by the reviewing team or subsequently identified as underpinning the poor performance of some records. These latter ones are covered under the reports for their respective data centres. Additionally, there were issues surrounding handling of PDC records due to the nature of their identifiers containing forward slashes which necessitated a specific work around within the DCS's underpinning Geonetworks service in order for these records to be rendered. Fortunately, this issue was resolved within the 4 week review window for this round of reviews and thus PDC records were able to be reviewed.

Overall comments

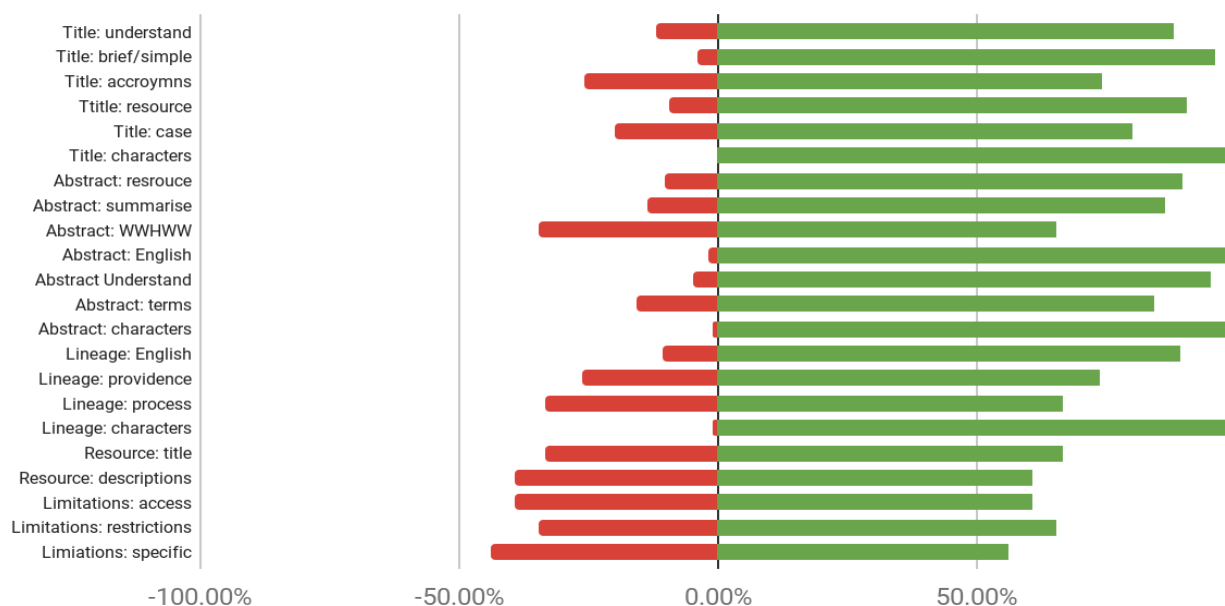
The following chart shows the percentage of records where the review question was either positive or negative. **These do not account for any selections of 'not applicable' or 'unsure'.** All questions are shown and though some may indicate areas that are of more concern than others it should be noted that not all questions carry equal weight in terms of their importance for record understandability. 2016 results are given too to aid comparison (though there are slight differences in the review questions).

October 2021 DCS Reviews: Summary Report

All reviewed records



All EDCs (2016)



Comparison of the results for the previous review round (note, abstract criteria order is slightly different) shows that overall the quality of the DCS records has improved in some areas and shown some decrease in others, but most remain largely in line. The variance between the this round of reviews and the 2016 results may be attributed to factors such as:

1. Simple variation in sampled records - especially that this sample has a lower representation from similar record types as seen in the 2016 reviews

2. Differences in reviewer pool and therefore variation in review responses - the reviewers did not go through any formal moderation process, though some discussions during the review process on how to assess did take place amongst the team.
3. Variation in review assessment criteria - especially for the latter part of the reviews looking at the access and usage limitations.
4. Shift in record presentation from GEMINI 2.1 records to GEMINI 2.3 and also the underpinning service running the data catalogue service and how it handles the rendering of content.

Areas where improvement can be noted include:

- Abstracts have improved slightly for the 'Who, What, When, Where and How' details of the dataset
- Legal constraints and access constraints - though this could also be down to improved handling of this information within the records as well as clearer assessment criteria. The rendering of these records was noted to be poorly done in the 2016 when there was also felt to be unclear assessment criteria and overall use of these fields by the reviewers. Efforts to improve this field were recommended and these results may be evidence of the impact of such improvements over recent years.

Areas where there appears to be a degrading of record quality are mainly associated with:

- Provenance details within the Lineage statement
- Process of data delivery details in the Lineage statement
- Resource descriptors

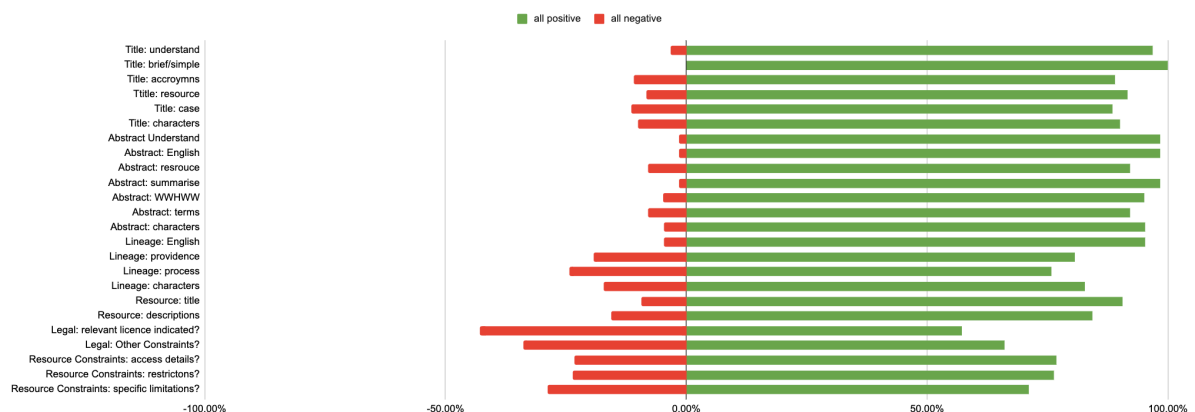
Individual Data Centre Results

Reviews for each data centre's records that were sampled will be made available to the data centre to allow further digestion and targeted actions to improve metadata quality where possible.

The following section provides summary reports for each data centre to show the variability between records from these different sources, highlighting variability in practice. 2016 charts have not been included here for space considerations - these are available in the 2016 report.

British Oceanographic Data Centre

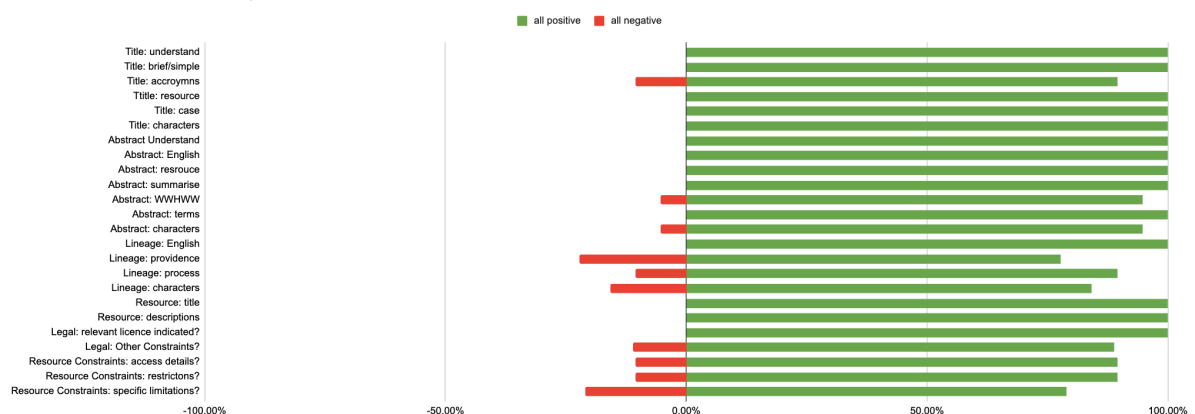
BODC 2021 - all sampled records



NOTES

Title and Abstract scores are slightly lower than the 2016 reviews with larger degradations apparently seen for lineage, legal constraints and access constraints fields. However, this may be attributed to the shift in sampling to reduce the prevalence of 'templated' records within the 2016 sample which were avoided in this sample. This is evident when the data are then split by non-collection datasets and 'collection' datasets, which show a much closer alignment to the 2016 results:

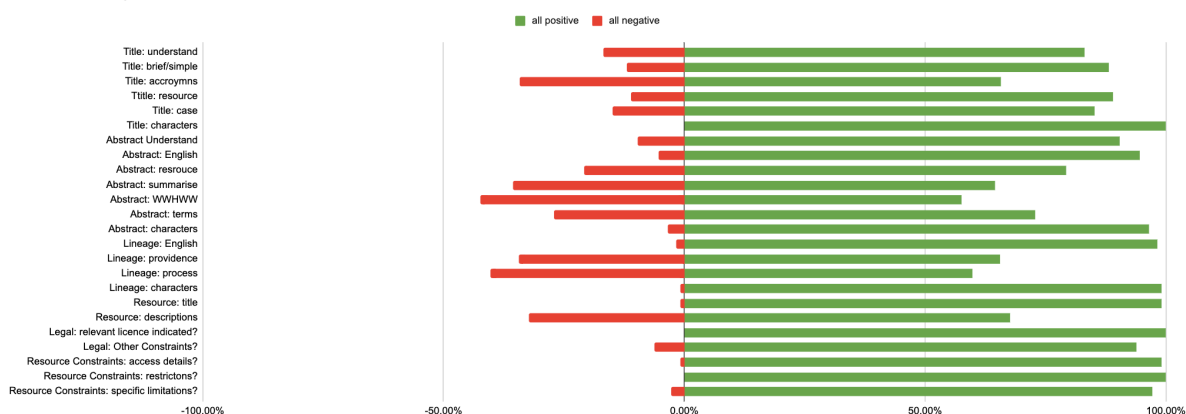
BODC - 'Collection' datasets only



This clearly demonstrates the advantages of metadata record templating for similar datasets and also the challenges when dealing with datasets that can not be easily templated. These account for the largest part of BODC records (855) generated indicating that, when these results are applied to all BODC records, overall BODC continues to perform well.

Centre for Environmental Data Analysis (CEDA)

CEDA - all sampled records 2021



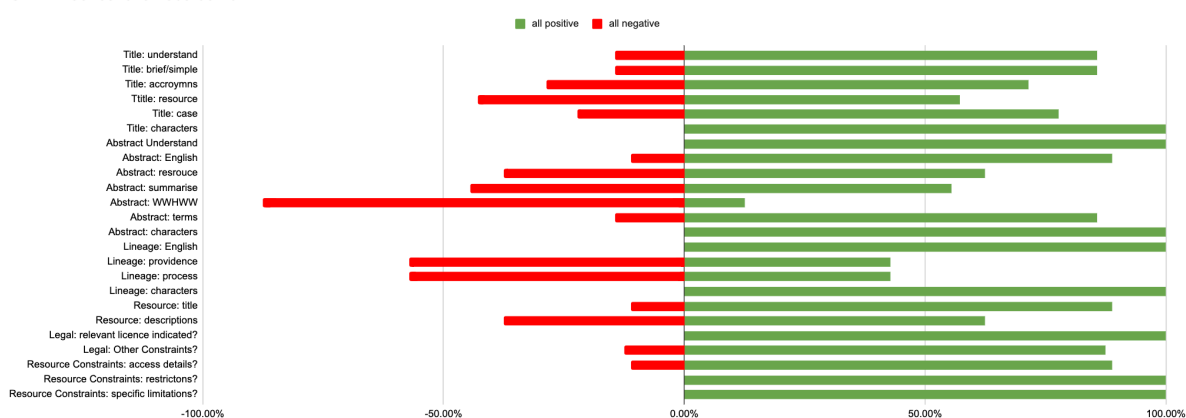
Notes

Compared to 2016 results, CEDA results have shown improvements in all areas, most notably for:

- Abstract summary of dataset contents
- Abstract 'who, what, when, where, how' information
- Resource information
- Access constraints information
- Legal constraints details

Given that CEDA records contained both Series and Dataset level records clarity is seen further when breaking down the results to look at Series level records and then datasets associated with collections of similar records:

CEDA - series level records 2021



Though 'series' record numbers are small overall, the small sample here should be factored in when examining these results. However, when compared to 2016 results for series level records an overall

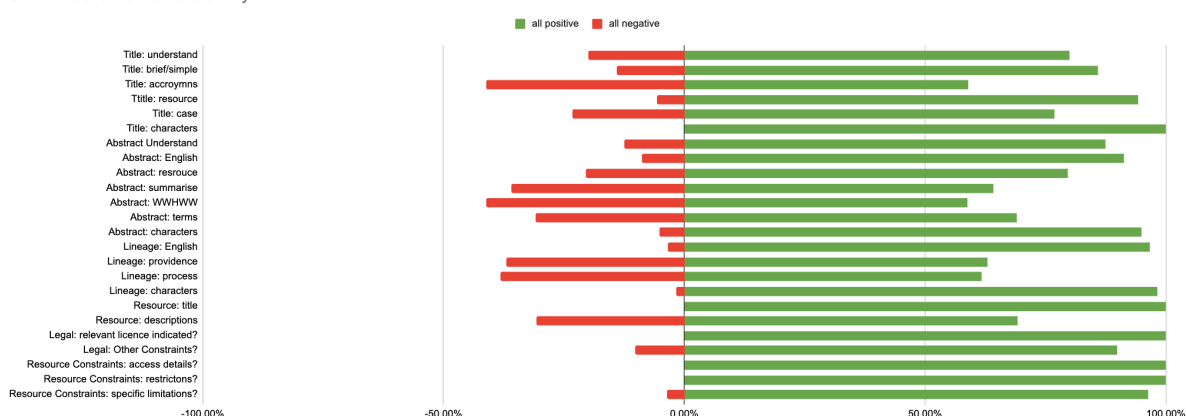
improvement is seen for all review criteria. Significantly improvements have been noted in the lineage, access constraint and usage limitation (legal) details - largely due to CEDA addressing these fields specifically when moving from GEMIN2.1 to GEMIN2.3. There has been a notable degradation in the Abstract 'understandability' which requires further review by CEDA and potentially remedial actions taken. However, as noted earlier, the low levels of records reviewed may simply show sampling issues here rather than overall issues with CEDA series level records. Similarly, the nature of these records to be more generic as they seek to act as umbrella records for underlying dataset records may indicate a need to adjust the review criteria for series level records.

Note, though CEDA do issue a handful of marked up 'non-geographic' datasets there small number were not covered by the sample. This is relatively new practice for CEDA and older 'non-geographic' datasets may still be issued as geographic datasets with a global bounding box. This is a known issue that CEDA will address in due course for these low-usage datasets.

Looking at those dataset records where templating was used for CEDA record, for which only sample records were reviewed (see first chart below) and comparing this with the remaining dataset level records shows that templating approach used within CEDA does work broadly well in producing good quality metadata records. There are areas where it has been shown to be better than non-templated approaches (e.g. abstract understandability), and others where there is some gain seen in the hand-crafted records (e.g. title acronyms and resource descriptions).

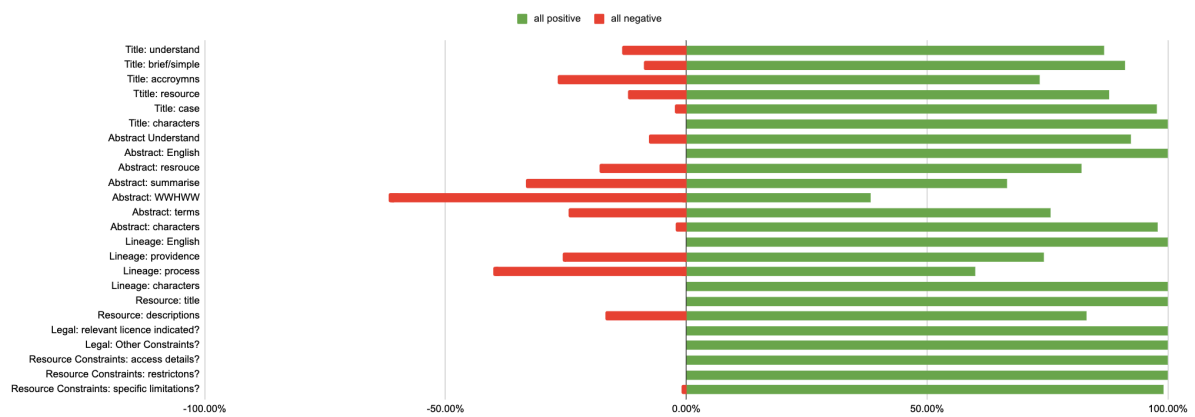
Given that CEDA make extensive use of such templating approaches to their records (accounting for nearly 3500 records alone of the 7500 records CEDA has produced to date - i.e. around half of their records), the impact of the templating approach is shown to allow great degree of scalability for CEDA's production of quality metadata records. It may also be possible for CEDA to make bulk changes to address any common, easily fixable issues with their records once full feedback has been reviewed to further improve metadata quality at scale through this approach.

CEDA - 'collection' datasets only 2021



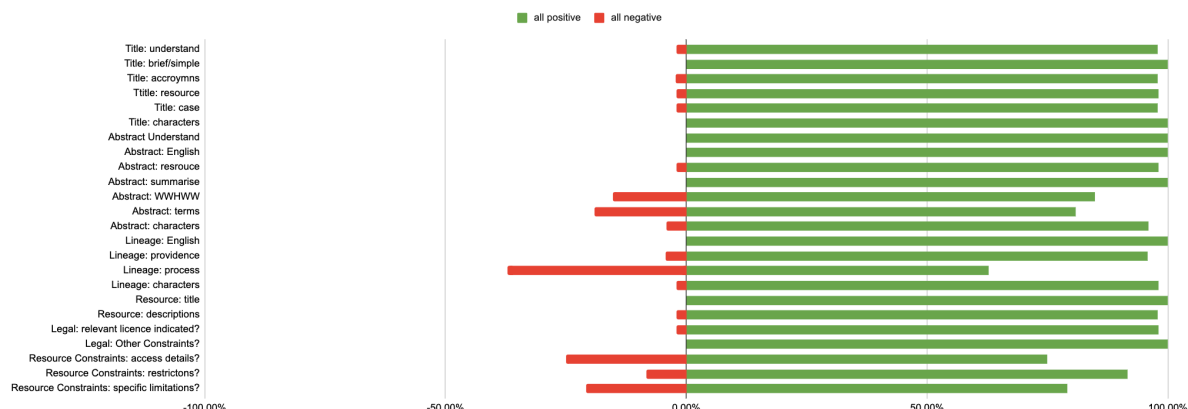
October 2021 DCS Reviews: Summary Report

CEDA: Non-collection dataset, 2021



Environmental Information Data Centre (EIDC)

EIDC - all reviewed records 2021

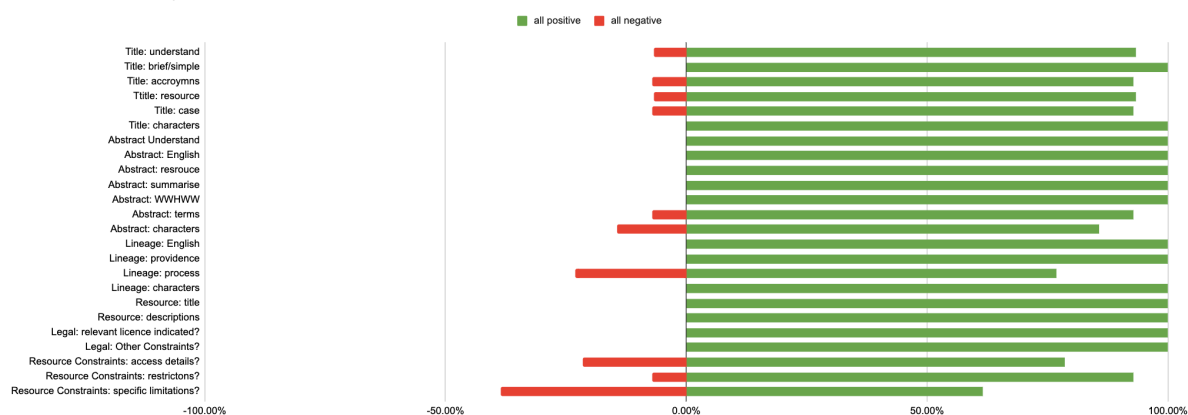


Notes

Compared with 2016 results EIDC records remain of very high quality in all areas with improvements seen in the title and abstract fields - particularly the 'Who, What, When, Where and How' abstract criterion. Lineage procedure details did, however, see a degradation from the previous review, whilst the resource constraints information also became an area of slight concern, though legal constraints (i.e. licensing information) improved. Taking these two fields together puts these broadly in line with the equivalent fields in the 2016 reviews.

When the samples of 'templated' dataset records were then examined on their own they were found to also be of overall high quality, though not quite as good as the non-collection dataset. In particular there were issues noted around resource constraint details:

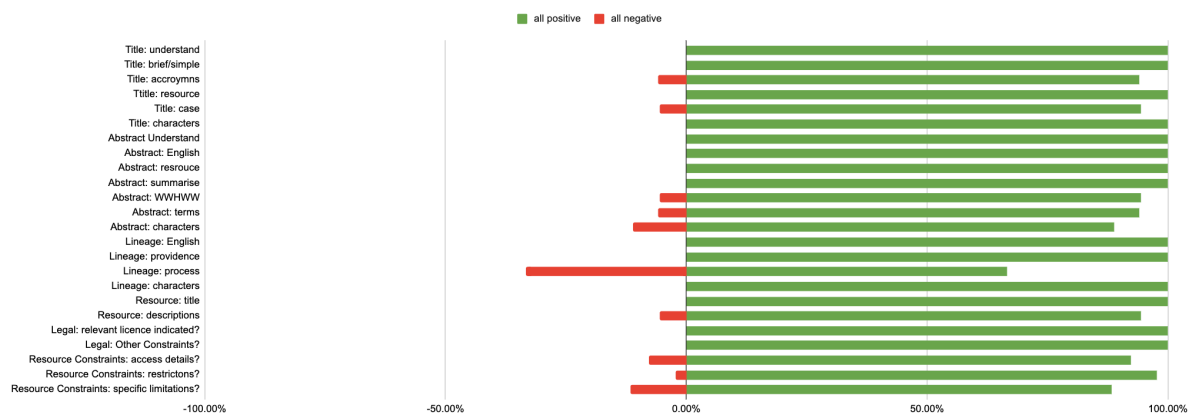
EIDC - collection sample dataset records 2021



October 2021 DCS Reviews: Summary Report

Finally, the handful (4) non-geographic EIDC records reviewed scored very well with full marks in over half the criteria, though lineage statements need some attention:

EIDC: non-geographic datasets, 2021



Polar Data Centre (PDC)

PDC: all reviewed records, 2021



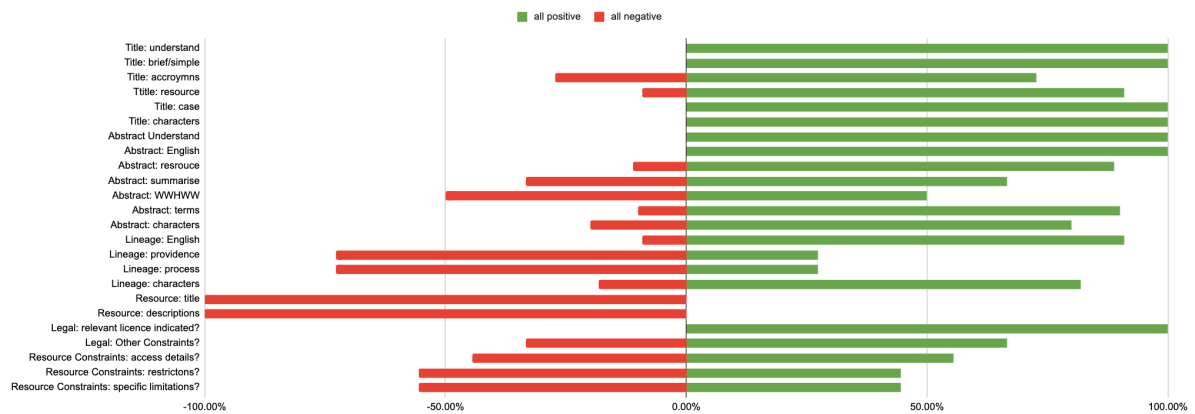
Notes

Since the 2016 review PDC records have shown improvements with regards to legal and access constraints. Most notably that specific licence information is now provided. However, no progress has been made with regards to resource information with just URLs being provided with no title or description as reported in 2016. Titles and Abstracts remain of high quality with some improvement seen within abstracts over the intervening period. Lineage details have shown some marginal improvement too, but largely remain an area of concern.

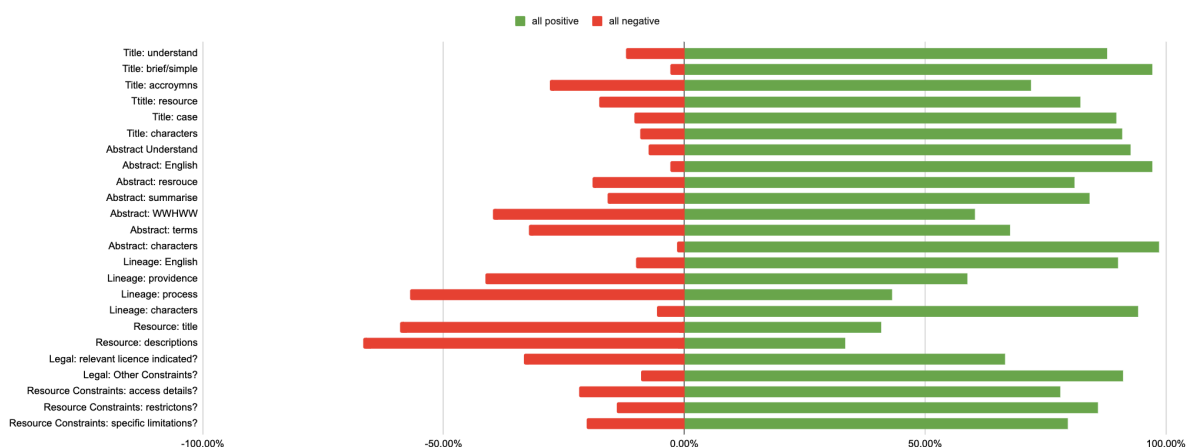
October 2021 DCS Reviews: Summary Report

Examining the 11 records taken as samples from ‘templated’ records indicates that there are some degradation in metadata quality through such templating approaches within the PDC across some review criteria:

PDC: collection sample datasets only , 2021



National Geophysical Data Centre (NGDC)

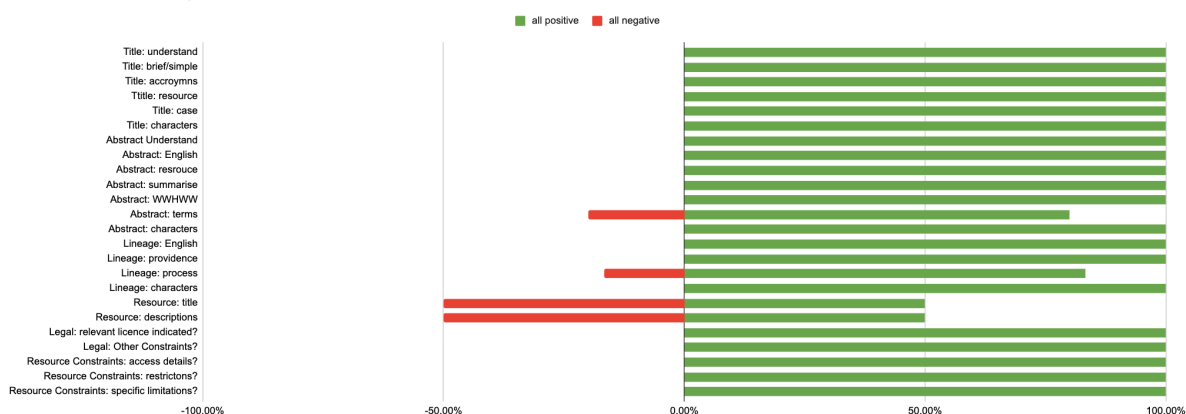


Notes

Since 2016 there have been some significant areas of improvement for NGDC records with regards to access constraints and usage limitation details, though reviewers did note that the information provided was not always tailored to the resource in question. There have also been some improvements in resource descriptions and titles. Abstract quality appears to have deteriorated, but this may be ascribed to the number of records where the reviewers felt that the items described were not actually data resources, but were other items such as presentations, reports etc and thus questionable if they ought to be listed within the NERC Data Catalogue Service as datasets/non-geographic datasets.

The NGDC samples also included 'templated' datasets. These numbers are small overall for NGDC and the related sample also small (just 6 records), but point to near perfect scores, with the main exception being resource title and descriptions.

NGDC - collections sample datasets



In contrast, the non-geographic datasets samples (28) were found to be of poor quality, mainly for the reason highlighted above that these were questionable resources to be added to the NERC Data Catalogue Service and therefore were marked down significantly within the reviews - most notably for the Abstract description's "resource" criterion - i.e. that in all cases reviewed the reviewers did not think that it described a data resource. Similarly, resource titles and descriptions and lineage information were found to be incorrect.

NGDC - non-geographic datasets



Metadata Quality Scores

A final step of the analysis was to score each record out of 100. This was calculated as a percentage of a combined score for the Major and Minor criteria (see appendix for scoring details) out of a possible score of 20 and then scaled to 100 to be comparable with the previous review rounds. Major and minor quality areas were originally established in review round 3 and have only been slightly adapted in this round to accommodate :

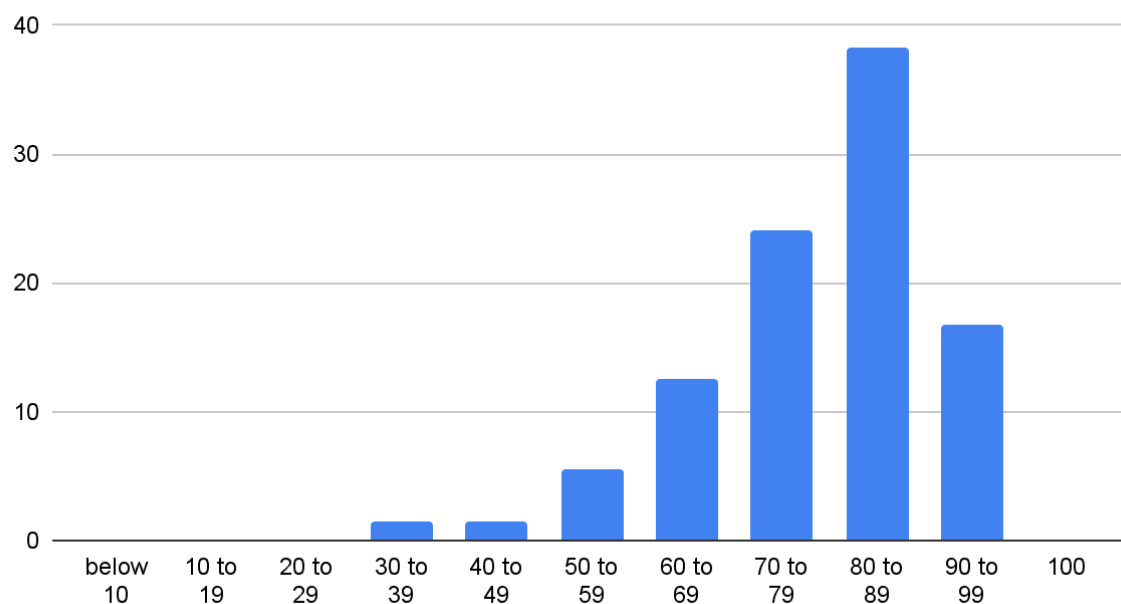
1. Splitting of the 'limitations' criteria into two distinct components: usage limitation and access constraints in this round
2. Introduction of 'Not Sure' and 'N/A' options - the former resulting in a ½ of the criterion's full mark if used.

A record failing all checks would score 0 and one passing all checks would score 100. This allows for easier comparison of metadata between the data centres themselves and with comparison between their scores in this round and earlier rounds when this was last undertaken to assess levels of improvements made.

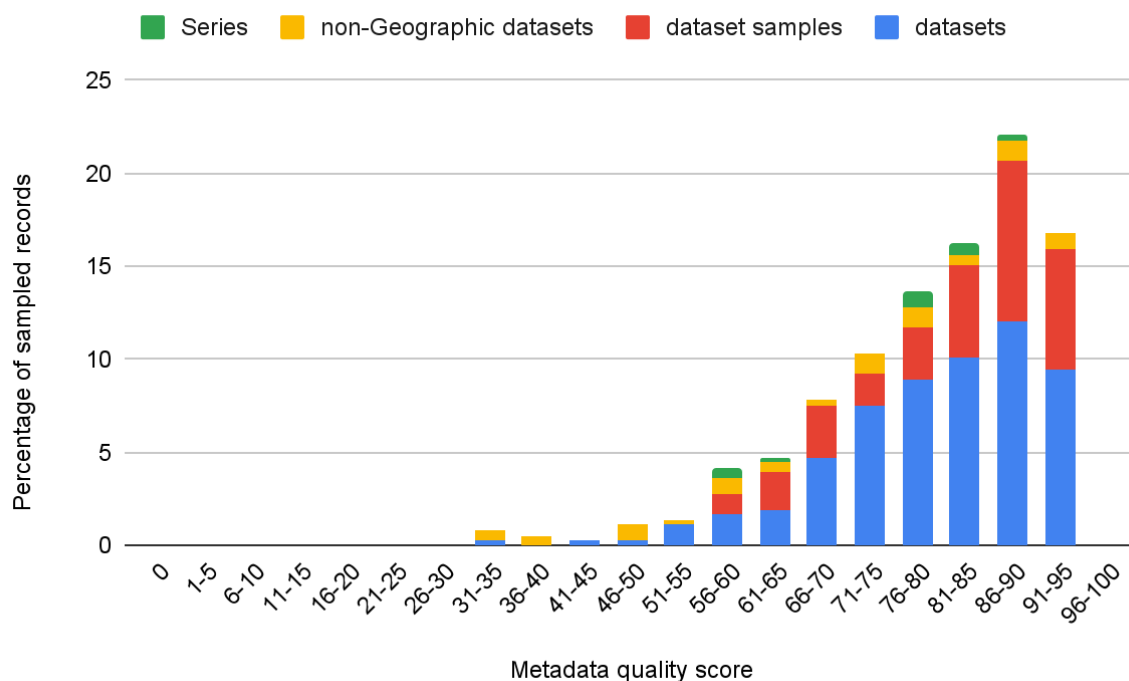
Overall Score

Overall the quality scores of records in the DCS records sampled are high with a peak score, when binned every 10%, between 80 and 90%. Compared to the 2016 result this suggests that there has been some convergence in overall scoring in the sampled records, probably due to the shift in records reviewed by data centre and similar sampling changes.

Overall record scores



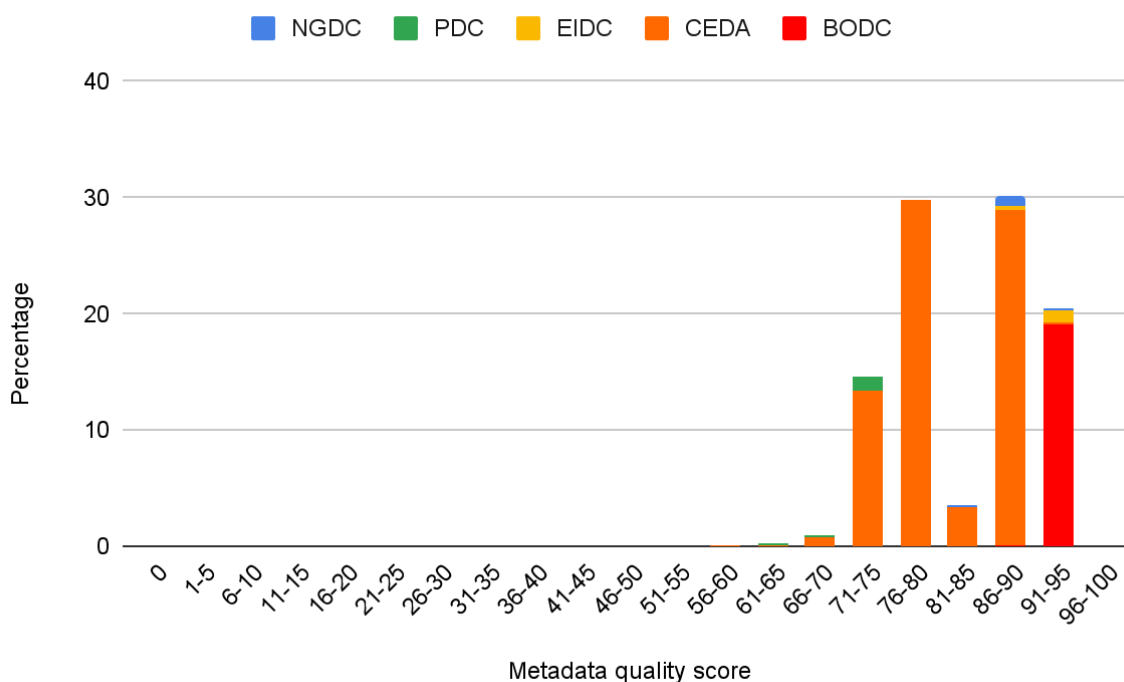
However, looking more closely at the scores such wide bins might actually be hiding some detail and this review round also sought to identify patterns between different record types. A re-binning to bins every 5% and then splitting the results by record type shows more detail:



Whilst the overall shape of the distribution remains similar to the 10% binned chart above, we can see that the main contribution to the lower scores comes from non-geographic datasets, whilst 'templated' datasets show greater variability than the pattern seen for other dataset records. Series records also show a wide spread in scores. However, the low numbers of non-geographic dataset

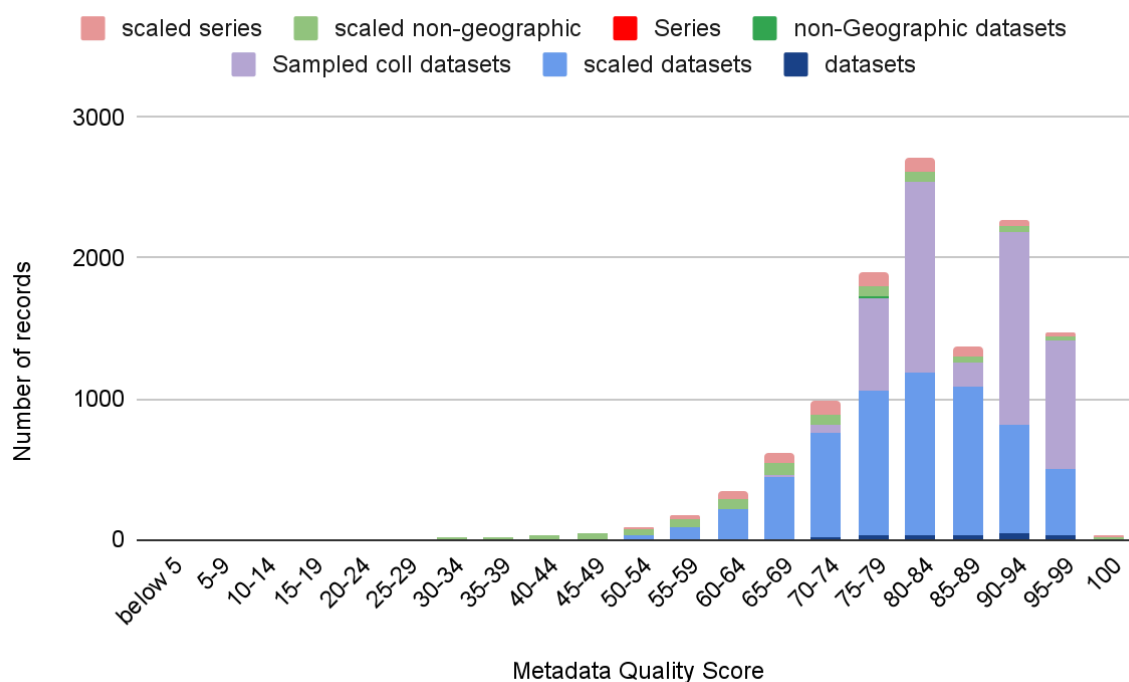
records and series level records reduce the applicability of these scores to all records of those types.

With the ‘templated’ datasets acting as examples for the other records within the collection these can then be applied to greater numbers of records within the NERC Data Catalogue Service. (Where more than one record for a collection was sampled an average score was taken). Scaling the scores for each collection by the average from the samples from the collection allows the results to be given for 4511 ‘templated’ records:



This clearly shows that overall templating leads to good to very good quality records. BODC, EIDC and NGDC have particularly benefited from this approach. CEDA makes heavy use of such record templating and record quality seems to fall into two distinct groups, related to specific collections. PDC tends to do less well through this approach. (Note also that the very high scores of BODC records and their over representation in previous rounds may well explain the higher scores seen in previous review rounds).

Taking scaled averages based on record type and data centre the results may be further scaled to the entire NERC Data Catalogue Service holdings to get an indicative average score of just over 80. Though results show a negative skew (i.e. modal peak is higher than mean value and long-tail towards lower numbers), using the mean and standard deviation for each record type within the sample and forming normalised distributions gives some indication as to the spread of results that would be seen:



This should only be taken as a very broad indicator, but does shed some light on the quality distribution by record type, and gives further context that templating approaches to metadata records provides a better outcome for dataset records on average. As this seeks to weight the scores by record type for the entire catalogue service contents this may provide a suitable benchmark for future review rounds to be compared against. **For this purpose, this round of reviews gives an overall scaled and weighted average for all data centres and record types metadata quality score of 80.3.**

The weighted overall average metadata quality score for each data centre is given below:

BODC	87.5
CEDA	80.4
EIDC	86.5
PDC	70.7
NGDC	72.8

I.e. BODC And EIDC are scoring above average, CEDA is around average and PDC and NGDC are below average. This is not surprising given that CEDA records contribute significantly to the overall contents of the NERC Data Catalogue Service.

Individual Data Centre Scores

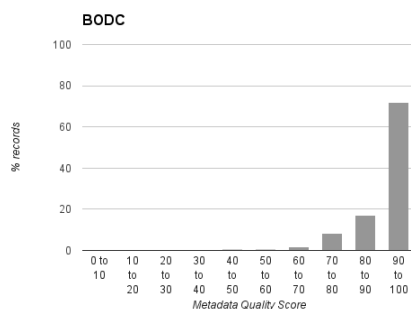
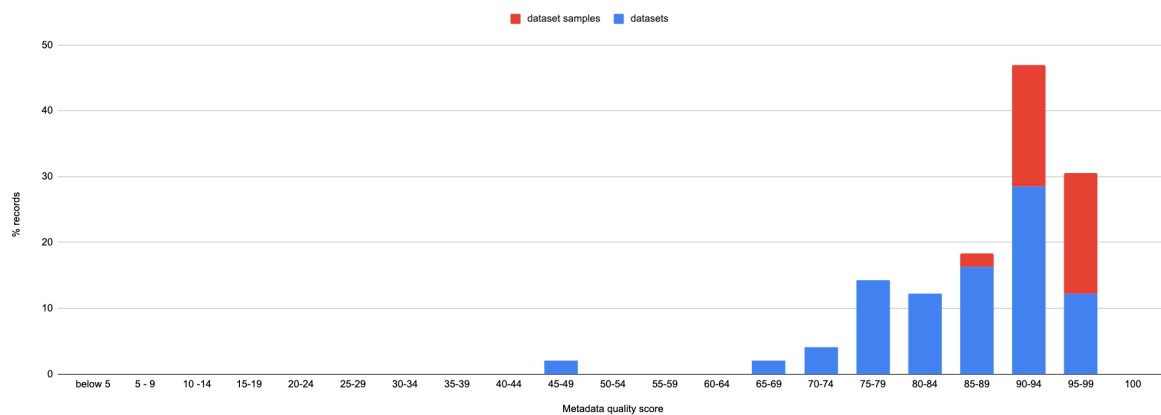
The following sections give the spread of metadata scores for each data centre compared with their 2016 results. The charts also show the breakdown by record type within the sample for each data

October 2021 DCS Reviews: Summary Report

centre, though the proviso of small numbers for the non-geographic and series level records should be borne in mind.

BODC

BODC Scores 202010

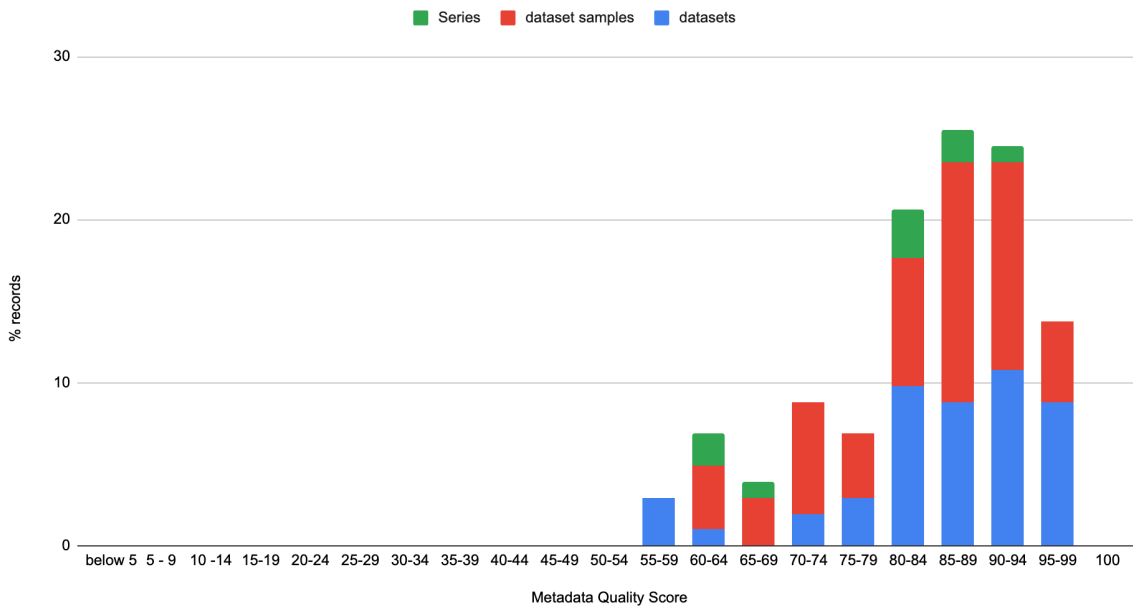


The BODC score remains high overall when compared to the 2016 spread (see inset chart to the left). However, though there is a marginal spread in the tail to lower scores in this round, which can be attributed to the reduction in the number of ‘templated’ datasets within the sampled BODC records. When templated datasets are examined the scores were found to be higher, as seen when the records are examined by type in the above chart, especially when the samples are then re-applied to

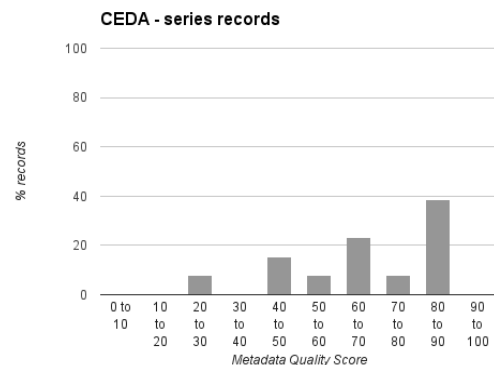
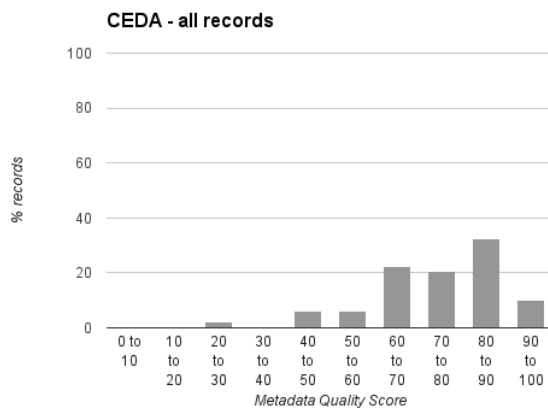
the collections that they represent.

CEDA

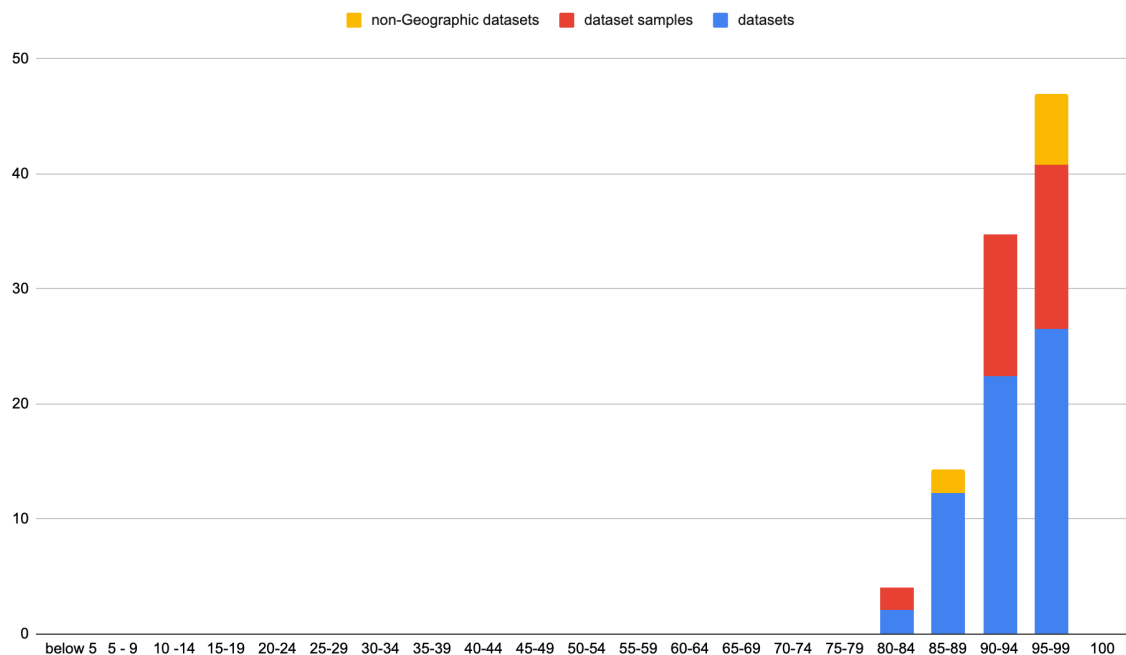
CEDA Scores 202010



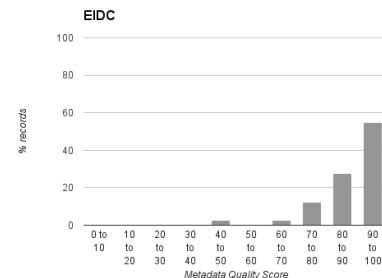
CEDA values show an overall increase in metadata quality since the 2016 scores (given below for all records and series level records), though the long tail towards lower scores (50 to 70) remains. Templating approached have served well for CEDA, especially with regards to some of their larger collections numbering thousands of records alone, but some templating shows some degradation in quality. This latter set of records can be identified not as templated datasets produced from standard scripts, but those where initial template datasets have been used from which other datasets are then 'cloned' and altered. This suggests that CEDA may need some care when taking this approach. Whether bulk improvements are easily possible here or not requires further checking by CEDA.



EIDC

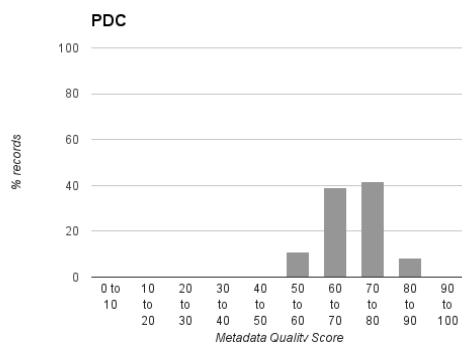
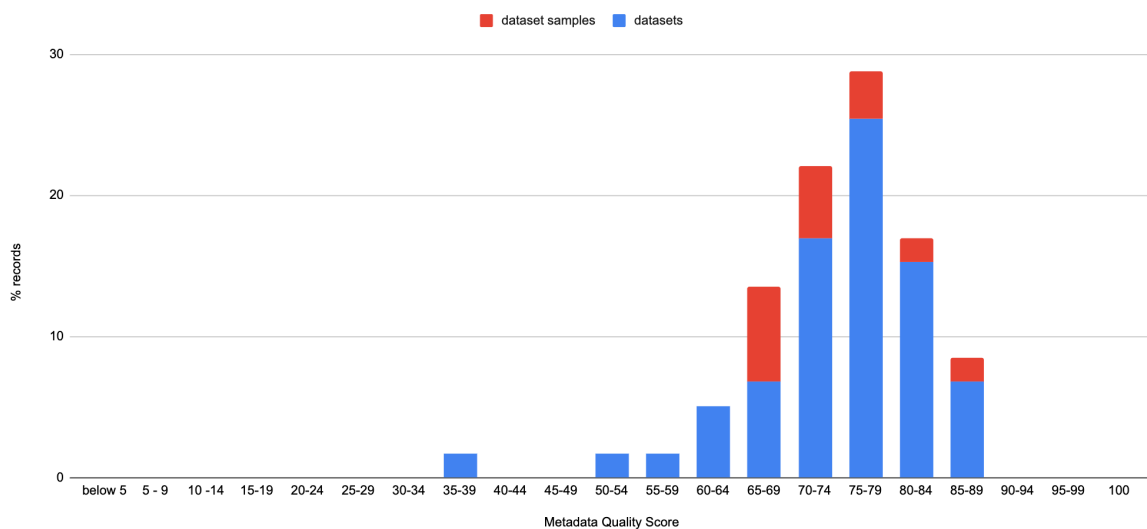


EIDC has made marked improvements in the quality of their metadata since the 2016 review (to the right) with many more records scoring 90 - 100. There are still some outliers requiring attention, but given that a doubling of EIDC records entering the DCS within the same period, this is an excellent result for EIDC. With a small exemption, both non-geographic and templated datasets also show a tendency to higher quality scores.



PDC

PDC Scores 202010



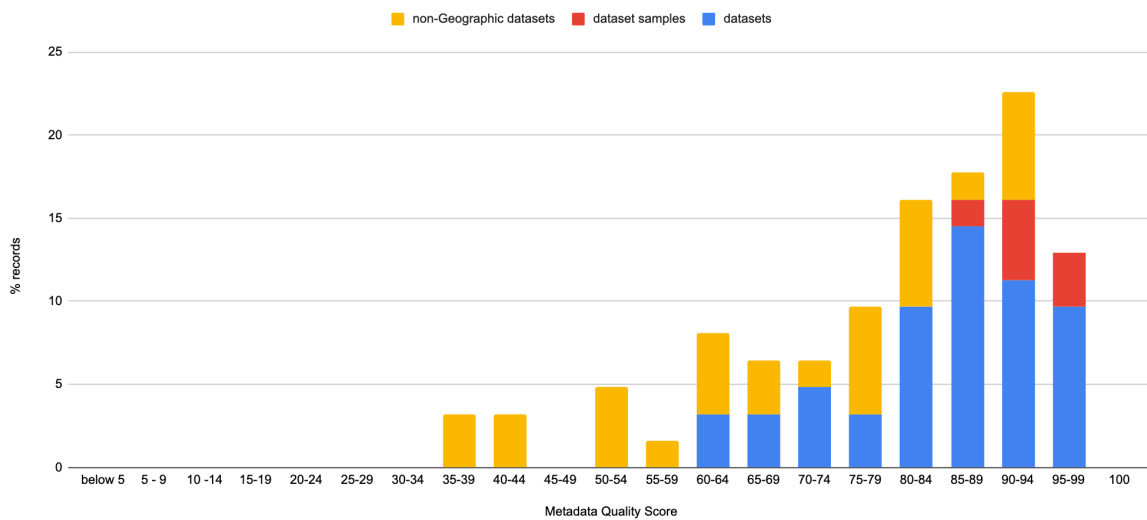
These figures show an overall shift in the distribution by around an extra 10 marks for all records over the 2016 results (left) demonstrating an overall improvement in Polar Data Centre records. This indicates that there have been widespread areas of improvement for PDC, though common areas of improvement for all records as noted earlier will help to move these scores higher.

Templated records appeared to fair slightly

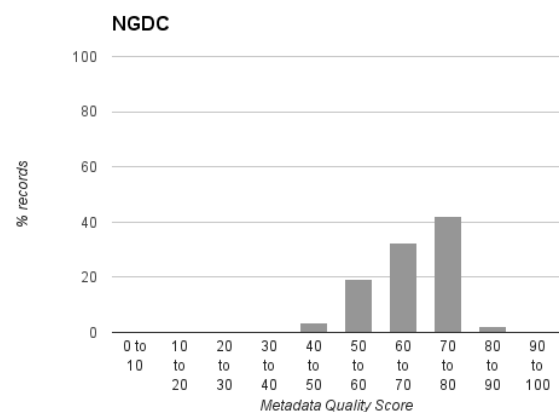
less well overall, though may provide an easy route for PDC to continue their improvements to metadata quality on a good proportion of their records.

NGDC

NGDC Scores 202010



NGDC scores have shown an overall shift to higher scores, but with a longer tail towards lower scores - especially within non-geographical datasets. Templated datasets tend to score higher on average for NGDC dataset level records.



Feedback from Reviews

During the latest round of reviews the review team had the opportunity to record comments about the three aspects: the review process itself; the metadata guidelines and the DCS service. Overall the responses from the reviewers was that the review system worked well with the only suggestion to alter the ordering of questions to match the order of record content to aid speed of finding content.

Reviews each took between 10 and 15 minutes each to complete, which was inline with expectations. With a requested ½ day to full day of effort per reviewer this would translate to roughly to between 20 and 40 reviews per reviewer. With one of the initial reviewer pool pulling out leaving

13 reviewers the number of reviews expected was between a minimum of 208 and maximum of 624, so the 358 was slightly lower than expected average of 390 reviews. However, effort per reviewer was quite variable with some reviewers exceeding this upper limit whilst others were much lower. Available effort by data centre also varied and this inevitably impacted on the reviews too. Ideally all data centres should be able to offer equal effort and all reviewers aim for similar levels of output in future rounds. Having established the time each review should take it should be possible to set a better benchmark for reviewers to aim for as a minimum in the future of 30 reviews each which should hopefully translate into sufficient effort being provided. However, the disparity in numbers of reviewer between different data centres is something that DOG should consider for future rounds. In this regard it would be worth noting that a number of reviewers specifically commented that it was useful for them to review records from other data centres as part of the process and such cross-centre working can help to propagate the best of metadata practises amongst the NERC data centres.

Use of Slack to aid the review process was appreciated by all reviewers and greatly assisted in resolving both technical issues and questions about the review process as it progressed. Additionally, it allowed a degree of friendly competition between reviewers spurring them on to undertaking more reviews and alleviating some of the monotony of the process.

Next steps and Recommendations

- Each data centre will receive access to the review responses allowing them to undertake analysis of errors/warnings noted for individual metadata records and reviewers' comments. Care should also be taken to address 'unsure' responses in the reviews too to strengthen the understandability further of those records.
- The appropriate data centre will decide what actions to take to amend the relevant records. Amending metadata is *not* a task of this project.
- Continue to make use of communication tools such as Slack to maintain connections between the data centres for future reviews/metadata questions.
- Following feedback from the review team and any comments that DOG wish to provide, adjustments will be made to the metadata quality process to address issues identified with the review criteria and methodology.
- Another round of checks will be scheduled as required by DOG - this should be spread over at least a week with remote sessions to initiate and end the review process with an opportunity to reconvene during the week as required.

Appendix A – Review Criteria

The following table shows the review criteria used along with the scores for answers given for Yes or Not Applicable and Not Sure/unsure, the latter scoring ½ of the full mark for a Yes/Not applicable answer. The last column shows the relative ratings given as per the 2016 reviews as to whether the criterion was classified as Major and Minor. Note, that whilst the mapping is mainly Major = 1 for yes or N/A this has been reduced to 0.5 for the final 5 marked criteria to address the difference with the 2016 criteria where the usage limitations and access constraints fields were previously reviewed as a combined field.

	yes or N/A	Not Sure/unsure	Understandability level
Title Field [Do you understand the title?]	1	0.5	Major
Title Field [Is the title brief and simple?]	1	0.5	Major
Title Field [Is the title free from unexplained technical terms or acronyms? (can be expanded in abstract)]	1	0.5	Major
Title Field [Does the title describe the resource rather than the activity/project which produced it?]	1	0.5	Major
Title Field [Is the title in sentence case?]	0.5	0.25	Minor
Title Field [Is the title free from html markup and non-utf8 characters? (For example: """)]	1	0.5	Minor
Common Issues (use 'other' to add additional comments)			N/A
Do you have any comments about the title?			N/A
Abstract Field [Do you understand the abstract?]	1	0.5	Major
Abstract Field [Is the abstract written in plain English?]	1	0.5	Major
Abstract Field [Does the abstract describe the resource rather than the activity/project which produced it?]	1	0.5	Major
Abstract Field [Do the first few sentences summarise the contents of the resource?]	1	0.5	Major
Abstract Field [Does the abstract explain Where, When, How, Why and Who?]	0.5	0.25	Minor
Abstract Field [Is the abstract free from unexplained technical terms or acronyms (could be defined in title)]	1	0.5	Major
Abstract Field [Is the abstract free from html markup and non-utf8 characters? (For example: """)]	1	0.5	Major
Common Issues (use 'other' to add additional comments)			N/A
Do you have any comments about the abstract?			N/A
[Is the lineage written in plain English?]	1	0.5	Major

[Does the lineage describe how the data resource came into existence?]	0.5	0.25	Minor
[Does the lineage describe the stages the resource has passed through before arriving at the data centre?]	0.5	0.25	Minor
[Is the lineage free from html markup and non-utf8 characters? (For example: """)]	1	0.5	Major
Common Issues (use 'other' to add additional comments)			N/A
Do you have any comments about the lineage?			N/A
[Do resource locators (download and links) include titles which are concise accurate accounts of the resource?]	1	0.5	Major
[Do resource locators (download and links) include descriptions which fully explain their purpose?]	0.5	0.25	Minor
Common Issues (use 'other' to add additional comments)			N/A
Do you have any comments about resource locators?			N/A
Legal Constraints [Does the Legal Constraint indicate the licence relevant to the resource?]	0.5	0.25	Major
Legal Constraints [Are other legal constraints listed?]	0.5	0.25	Major
Common Issues (use 'other' to add additional comments)	0.5	0.25	N/A
Do you have any comments about Legal Constraints?	0.5	0.25	N/A
Resource Constraints [Are details given on gaining access?]	0.5	0.25	Major
Resource Constraints [Are there details on access restrictions?]	0.5	0.25	Major
Resource Constraints [Are limitations on public access and access constraints specific to the particular resource?]	0.5	0.25	Major
Common Issues (use 'other' to add additional comments)			N/A
Do you have any comments about Resource Constraints?			N/A
If you'd like to make any general comments about this metadata record, you can do so here.			N/A