

# Collections development selection and appraisal criteria



---

15 March 2022

**Public**

Copyright © 2022 University of Essex. Created by UK Data Archive, UK Data Service.

Version: 07.00



---

## Table of contents

<b>1. Scope .....</b>	<b>3</b>
<b>2. Selection and appraisal strategic values, priorities, and criteria .....</b>	<b>3</b>
2.1 Strategic values .....	3
2.2 Stakeholder priorities .....	3
2.3 Appraisal criteria .....	3
2.4 Out of scope materials.....	4
<b>3. Curation categories.....</b>	<b>5</b>
<b>4. Appraisal criteria for submissions and data discovery .....</b>	<b>5</b>
4.1 Stage 1: Criteria relating to strategic high-level value.....	5
4.2 Stage 2: Criteria relating to user need and analytic value .....	6
4.3 Stage 3: Criteria relating to potential usability and accessibility.....	6
<b>5. UK Data Service's Appraisal Grid .....</b>	<b>6</b>
5.1 Appraisal Grid scope .....	6
5.2 Presenting submission to UK Data Service Data Appraisal Group.....	7
5.3 Making decisions for ingest and access routes.....	7
<b>Appendix A: UK Data Service Collection Development Appraisal Grid.....</b>	<b>8</b>

## 1. Scope

This document sets out criteria for selecting and appraising data collections across the UK Data Service.

A coordinated and robust approach to appraisal is required to ensure that:

- Acquisition decisions will be based on explicit priorities and criteria which can be justified.
- Ingest activities can be prioritised and data collections will follow an appropriate ingest and access pathway.
- Business metrics can be conducted on Collections Development activities.

## 2. Selection and appraisal strategic values, priorities, and criteria

### 2.1 Strategic values

The UK Data Service's Collection Development Policy (CDP) states that it acquires data to meet three strategic values: potential secondary use and analysis for research, teaching and learning use, and replication and validation of research.

### 2.2 Stakeholder priorities

These strategic values align with data collections driven by research and policy needs such as:

- longitudinal data on people and families
- addressing multidisciplinary challenges (e.g. biosocial data, links with environmental data)
- data about organisations
- customer databases
- social media as research data
- cross-sectional data on people and families
- international data access
- administrative data
- new and novel forms of data.

### 2.3 Appraisal criteria

The criteria consider the following:

---

- 
- **Relevance to the remit:** The materials meet the UKRI ESRC's current strategies and are within the 'Scope of Collections'. This includes being high-quality, authoritative, reliable and up-to-date with 'good temporal and spatial coverage', with long and consistent time series (as appropriate), and strong opportunities for comparable research.
  - **Scientific or historical value:** The materials are considered important resources either for research purposes and/or meet current demands in research trends and/or there is research and/or teaching activity in the subject area covered by the data.
  - **New sources or types of data:** The materials meet emerging research areas within remit and/or have been requested by users.
  - **Uniqueness/risk of loss:** The materials form a rare and complete resource of value to the research community and/or are at risk of loss if not acquired; data collections which are costly to reproduce or to which continued access to data is threatened; data are not generally available in any other form e.g. paper.
  - **Usability/redistribution/operational benefits:** The materials have been well prepared according to data format/documentation and ethical/legal guidelines and fall within Scope; accession to the Service makes the resources more accessible. Where data are made completely open (publicly accessible without need for registration) by the data owners, the UK Data Service may continue to make accessible versions of the data because it fits well with the portfolio of other data<sup>1</sup>.
  - **Reproducibility of existing data outputs:** The materials are required to reproduce published or unpublished research, and/or undertake a restudy.

## 2.4 Out of scope materials

Where materials fall outside of scope, efforts are made to determine whether another archive, data centre or institutional repository is more appropriate for dissemination, curation or preservation of these data.

There are also certain criteria for not accepting data even though the value is deemed to be high:

- Legal and ethical considerations such as consent, copyright and/or Data Protection issues which are unable to be satisfactorily resolved. In these cases, full use of the data would not be possible without infringing legislation.
- Lack of sufficient contextual materials to enable re-use.

---

<sup>1</sup> Anticipated access method will not affect the appraisal of data.

- Old/unreadable formats that are difficult to convert or make usable, or impossible to recover.

### 3. Curation categories

The CDP lays out five possible classifications for data collections which correspond to the five curation categories below:

**CURCAT1:** Data collections selected for long-term curation.

**CURCAT2:** Data collections selected for short-term management.

**CURCAT3:** Data collections selected for ‘delivery’ only.

**CURCAT4:** Data collections selected for “discovery” only.

**CURCAT5:** Data collections selected for preservation only (note: this category falls outside the scope of the UK Data Service Appraisal, and is handled by the UK Data Archive).

### 4. Appraisal criteria for submissions and data discovery

A set of criteria are used for judging strategic high-level value, assessing user need, analytic value, usability, and accessibility. These are classified into three stages. Value criterion is used as a first stage for appraising data for selection. If data do not pass the first stage, if they are out-of-scope and do not meet the Collections Development remit, they are unlikely to pass to Stage 3 to be assessed for their accessibility and/or usability. Data arising from UKRI ESRC awards (CURCAT 2) are not appraised using value criterion and are typically subject to Stage 3 checks only.

#### 4.1 Stage 1: Criteria relating to strategic high-level value

- UK Data Service remit as set out by the CDP strategic values [section 1.1] (at least one value must be met).
- Data relevant to stakeholder research priorities (which may change over time) [section 1.2].
- Collections with a national mandate to share (e.g. UKRI ESRC Research Data Policy, UKRI ESRC-funded longitudinal studies).
- New and novel forms of data from emerging research areas acquired to meet researcher needs and/or selected for acquisition by stakeholders.
- Data from key long-term series already held by the UK Data Service with existing high usage (mostly produced by government).

## 4.2 Stage 2: Criteria relating to user need and analytic value

- Data useful for supporting decision-making and informing policy.
- Data useful for international, longitudinal or comparative research (e.g. time series or harmonised).
- Authoritative source of high quality data that are/have been widely used and cited.
- Data directly linked to a peer-reviewed published paper.
- Microdata otherwise not available to the research community (e.g. disclosive data).
- Data directly requested by a critical number of users.
- Unique, unrepeatable data and those that are costly to reproduce.

## 4.3 Stage 3: Criteria relating to potential usability and accessibility

- Data that have definable conditions within the access spectrum.
- Legal and ethical agreements and protocols used do not prohibit re-use, e.g. consent to share/reuse data, IPR, copyright, UK GDPR.
- Data have appropriate supporting documentation/contextual information.
- Data are produced in, or can be converted at minimal or no cost to, a suitable format for reuse through a UK Data Service delivery platform (without need for specialised software).
- Third party host is a reliable and reputable supplier of data.
- Third party host can confirm that data supplied are likely to be stable and not change over time.
- Third party supplier/host can confirm that suitable preservation processes are used to manage and curate data.
- Third party host/supplier can confirm that a persistent identifier can be made available.

# 5. UK Data Service's Appraisal Grid

## 5.1 Appraisal Grid scope

The UK Data Service uses an Appraisal Grid (Appendix A) for the transparent implementation of its Collections Development Policy. The grid is included in the Pre-ingest application of the UK Data Service's dissemination/access system Endeavour.

---

## 5.2 Presenting submission to UK Data Service Data Appraisal Group

New data offers are initiated by depositors via their account. A member of the Collections Development team will first appraise the offer by completing the appraisal grid.

For an existing data series, a light touch appraisal undertaken. For a new series, the completed grid is discussed by Collections Development and may be taken forward to the UK Data Service Data Appraisal Group (DAG) for further discussion.

At least one of the UK Data Service high-level value criteria must be checked for the offer to move forward to the DAG.

## 5.3 Making decisions for ingest and access routes

The proposer completes the appraisal grid (Appendix A) with recommendations for ingest and access pathways. The DAG uses a collection's completed Appraisal Grid, together with the presenter's case for support, to agree decisions about the collection's outcome. DAG makes decisions about:

- Whether to acquire the collection.
- Its curation category (CURCAT 1-5).
- Its anticipated access and delivery mechanism; options for CURCAT 1 and 3 are:
  - Access type: Open; Safeguarded (End User Licence; Special Licence; Additional Conditions); Controlled (approved/accredited researcher access).
  - Data delivery system: Download system; Nesstar; UKDS.Stat; Infuse, WICID; Casweb; SecureLab remote access (including References Drive and researcher-imported data; SecureLab Safe room; approved Safe Pod(s); QualiBank; ReShare; AWS Platform; Open API; Other added value.
  - Suitability for developing into a teaching dataset.
- Whether it might merit any high-level promotion, user-oriented activity such as a guide, workshop or any special communication activity.
- A possible or expected release date; how urgent ingest might be, deciding low, medium or high priority for processing.

All decisions are recorded in the appraisal grid section of the Pre-Ingest Application in Endeavour.

## Appendix A: UK Data Service Collection Development Appraisal Grid

	Yes	No	Comment
<b>Strategic value</b>			
Potential secondary use and analysis for research			
Teaching and learning use			
Replication and validation of research			
<b>Relevance to the UKRI/ESRC Data Strategy</b>			
Longitudinal data on people and families			
Addressing multidisciplinary challenges (e.g. biosocial data)			
Data about organisations			
Consumer databases			
Social media as research data			
Cross-sectional data on people and families			
International data			
Administrative data			
New and novel data			
Does the collection have a mandate to share with UK Data Service?			
Have these data been recommended by UK Data Service's Advisory Board?			
<b>User need and analytic value</b>			
Data useful for supporting decision-making and policy formation			
Data useful for international, longitudinal or comparative research (e.g. time series or harmonised)			
Authoritative source of high quality data and/or widely cited			



	Yes	No	Comment
Microdata otherwise not available to the research community (e.g. disclosive)			
Data directly requested by a critical number of users			
Unique, unrepeatable data or data that are costly to reproduce			
Are there related data already in the Collection?			
Statement on potential use:			
<b>Usability and accessibility</b>			
Data have definable conditions within the access spectrum			
Legal and ethical agreements do not prohibit re-use, e.g. consent to share/reuse data, IPR, copyright, DPA			
Data has adequate supporting documentation/contextual information			
Data are produced in, or can be converted at minimal or no cost to, a suitable format for reuse (without need for specialised software)			
Do data exceed 5GB in size?			If yes, inform IT teams
Third party host (e.g. API): is reliable and reputable			
Third party host: is the data likely to change/be updated over time?			
Third party host: suitable preservation processes are used to manage and curate data			

	Yes	No	Comment
Third party host: a persistent identifier can be made available (robust access)			

<b>Proposer Recommendations</b>			
Proposer name (autocomplete)			
Recommended acceptance outcome?	<input type="checkbox"/> Accept <input type="checkbox"/> Unable to accept	<input type="checkbox"/> Referred	
Curation Category	<input type="checkbox"/> CURCAT 1 (long-term curation) <input type="checkbox"/> CURCAT 2 (short-term management)	<input type="checkbox"/> CURCAT 3 (delivery only) <input type="checkbox"/> CURCAT 4 (discovery only) <input type="checkbox"/> CURCAT 5 (preservation only)	
Anticipated access (CURCAT 1 and 3)	<input type="checkbox"/> Open, without registration <input type="checkbox"/> Safeguarded (End user licence) <input type="checkbox"/> Special conditions	<input type="checkbox"/> Special licence <input type="checkbox"/> Controlled access (Approved researcher)	
Data Delivery	<input type="checkbox"/> Download System <input type="checkbox"/> Nesstar <input type="checkbox"/> UKDS.Stat <input type="checkbox"/> InFuse <input type="checkbox"/> WICID <input type="checkbox"/> Casweb <input type="checkbox"/> SecureLab remote access	<input type="checkbox"/> SecureLab safe room <input type="checkbox"/> QualiBank <input type="checkbox"/> ReShare <input type="checkbox"/> Open API <input type="checkbox"/> Other added value	
Suitable for high-level promotion/workshop/communication activity?			
Release date deadline			
Expected release date			
Recommended Priority Level	<input type="checkbox"/> High	<input type="checkbox"/> Medium	<input type="checkbox"/> Low

[www.ukdataservice.ac.uk](http://www.ukdataservice.ac.uk)

[help@ukdataservice.ac.uk](mailto:help@ukdataservice.ac.uk)

We are supported by the Universities of Essex, Manchester, Southampton, Edinburgh, University College London and Jisc. We are funded by UKRI through the Economic and Social Research Council.