Collections development selection and appraisal criteria

Table of contents

1. Scope

This document supports the Collections Development Policy (CDP) and sets out and implements the criteria for selecting and appraising data collections for inclusion in the UK Data Service collection. This approach ensures that:

- Acquisition decisions will be based on explicit priorities and criteria which can be justified, are transparent and accountable.

- Ingest activities can be prioritised and data collections follow an appropriate ingest and access pathway.

- Business metrics can be conducted on Collections Development activities.

2. 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, 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 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.

2.1 Stage 1: Criteria relating to strategic high-level value

- UK Data Service remit as set out by the CDP strategic values [section 4.2] (at least one value must be met).

- Data relevant to stakeholder research priorities (which may change over time) [section 4.1].

- 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).

2.2 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.

2.3 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.

3. UK Data Service's Appraisal Grid

3.1 Appraisal Grid scope

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

Appendix A: UK Data Service Collection Development Appraisal Grid

	Yes	No	Comment
Strategic value			
Potential secondary use and analysis for research			
Teaching and learning use			
Replication and validation of research			
Relevance to the UKRI/ESRC Data Strategy			
Area and development studies			
Demography			
Economic and social history			
Economics			
Education			
Environmental planning			
Human geography			
International relations			
Linguistics			
Management and business studies			
Politics			
Psychology			
Science and technology studies			
Social anthropology			
Social policy			
Social statistics, methods and computing			
Social work			
Socio legal studies			
Does the collection have a mandate to share with UK Data Service?			

	Yes	No	Comment
Have these data been recommended by UK Data Service's Advisory Board?			
User need and analytic value			
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			
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:			
Usability and accessibility			
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			

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

Proposer Recommendations		
Proposer name (autocomplete)		
Recommended acceptance outcome?	□ Accept □ Unable to accept	□ Referred
Curation Category	□ CURCAT 1 (long-term curation) □ CURCAT 2 (short-term management)	□ CURCAT 3 (delivery only) □ CURCAT 4 (discovery only) □ CURCAT 5 (preservation only)
Anticipated access (CURCAT 1 and 3)	□ Open, without registration □ Safeguarded (End user licence) □ Special conditions	□ Special licence □ Controlled access (Approved researcher)
Data Delivery	□ Download System □ Nesstar □ UKDS.Stat □ InFuse	□ SecureLab safe room □ QualiBank □ ReShare □ Open API

Proposer Recommendations			
	□ WICID		□ Other added value
	□ Casweb		
	□ SecureLab remote access		
Suitable for high-level promotion/workshop/communication activity?			
Release date deadline			
Expected release date			
Recommended Priority Level	□ High	□ Medium	□ Low

www.ukdataservice.ac.uk

help@ukdataservice.ac.uk