Annex 1: The ADR UK Flagship Data Framework

Introduction to the ADR UK Flagship Dataset Framework

ADR UK flagship datasets are a curated selection of research ready administrative datasets for which ADR UK coordinates activity and provides targeted investment to maximise their research use and public value. Research ready data refers to administrative data that has been sufficiently prepared, documented and structured to enable meaningful, ethical and secure research.

The concept of flagship datasets is grounded in ADR UK's evolving understanding of what constitutes research ready data. This includes insights from the February 2022 roundtable on "What is research ready data?", which drew on a systematic review of the published literature (McGrath-Lone et al., 2022), as well as the Data Utility Framework adopted by HDR UK (Gordon et al., 2021).

The purpose of the Flagship Dataset Framework is to increase the research use and public good of linked administrative data through prioritised and sustained investment in a strategic set of datasets. Through this approach, ADR UK aims to build trust, raise awareness and streamline the researcher journey, supporting reproducible research that contributes to improved social and economic outcomes.

The Framework sets out principles for identifying and curating research assets with the greatest potential value to society. It has been endorsed by the ADR UK partnership as an internal working document and will be refined for publication in due course.

What is the ADR UK Flagship Dataset Framework?

There are six guiding principles for the ADR UK Flagship Dataset Framework, the first five are directly informed by the systematic review of what makes administrative data research ready by Louise McGrath-Lone et al (2022). The sixth principle demonstrates our commitment to supporting researchers to use these complex data in the public good.

In the following tables each principle is broken down to describe the ADR UK flagship dataset minimum standards and the expectations for maintaining and enhancing flagship datasets as enduring research assets over the course of the ADR UK's investment period 2026-2031

The ADR UK Flagship Dataset Framework

Guiding Principle	Guiding principle description	Minimum standard: evidence for self-assessment	Enduring research asset requirements
ACCESSIBLE	ADR UK flagship datasets must be findable in a catalogue and available for use by all accredited researchers via an ADR UK trusted research environment with approved legal gateways for research (e.g., Digital Economic Act (DEA), GDPR).	i. Link to catalogue entry within one of ADR UK's trusted research environments available for all accredited researchers. ii. A clear mechanism in the documentation for resolving researcher feedback on any **TRE related data access** issues including output clearance and for feedback to inform service improvements throughout the researcher journey. iii. A commitment to work towards the enduring research asset requirements over the investment period 2026-2031 including a nominated individual or organisation responsible for coordinating this guiding principle.	Information should be made publicly available about the timeframes involved in accessing each dataset including getting outputs approved - ideally taking account of the governing process within the control of the trusted research environment and those dependent on researchers (Five Safes). • The **first step is to ensure transparency of information** about access timeframes to support researcher planning. • The **longer-term ambition should be to reduce any unnecessary delays** in the research access journey – streamlining this journey over time. • **Guidance, standards and best practice** can be developed by the partnership in due course including whether there are any circumstances where datasets not able to provide access within a maximum timeframe should be eligible for flagship status. Identify and resolve systemic TRE related dataset access issues to continuously smooth the researcher journey. Communicate service improvements to all current and future users.
BROAD	An ADR UK flagship dataset must be an administrative dataset with national coverage across at least one UK nation. It must be	i. Description of the component data sources including which are administrative data and how they are linked or linkable to at least one other dataset. ii. A 1–2-page strategic case for publication	**Promotion & awareness raising:** To widen the research use of flagship datasets and capitalise on their broad appeal - activity should be developed to raise awareness of these strategically important data and their potential.

Guiding Principle	Guiding principle description	Minimum standard: evidence for self-assessment	Enduring research asset requirements
	linked or linkable to at least one other dataset to enable strategically important or cross-cutting research in the public good.	iii. A commitment to work towards the enduring research asset requirements over the investment period 2026-2031 including a nominated individual or organisation responsible for coordinating this guiding principle.	**Research funding:** In addition, flagship datasets should be promoted alongside relevant funding opportunities. This could be via the ADR UK Research Fellowship or PhD funding opportunities as well by generating additional research co-funding or identifying and other relevant research opportunities. Funding opportunities can also be signposted to via the ADR UK Funding Opportunities webpage: Funding opportunities - ADR UK and promoted through our regular ADR UK channels including newsletters. **Research agenda development and promotion:** Further work to establish and promote the research potential of the data e.g., research agenda development via stakeholder or public engagement
CURATED	ADR UK flagship dataset quality must be transparently communicated and managed so researchers can make informed decisions on its use. This would include a formal feedback loop between researchers to those involved in creating the dataset to iteratively improve the quality of the data and documentation and therefore the resultant research outputs.	i. A short statement outlining the **mechanisms to capture and respond to feedback** from data users around the quality of the data (and associated documentation) ii. The **quality of the data must be documented** and meet the ADR UK minimum standards of documentation, including: • Information about the quality of the component datasets • Linkage quality and linkage rates • Strengths, limitations and constraints of the data iii. A commitment to work towards the enduring research asset requirements over the investment period 2026-2031 including a	Strive **towards continuous improvement of the quality of the data** documentation and where possible the underlying data: • Regularly **review feedback and identify issues** around the quality of the data and documentation received by data users and source data owners. • **Explore and resolve quality issues** identified with the data and/or documentation including the potential impacts on users. • Regularly **review and update documentation** to communicate quality issues and involving all relevant organisations involved in creating the flagship dataset where applicable/possible (see DOCUMENTED).

Guiding Principle	Guiding principle description	Minimum standard: evidence for self-assessment	Enduring research asset requirements
		nominated individual or organisation responsible for coordinating this guiding principle. NOTE: the minimum standards of documentation are due for review by March 2026 and will be revised to incorporate minimum standards for documenting the quality of flagship datasets.	• Develop a **shared understanding about the data** by drawing on the collective expertise and experience of data suppliers, data users e.g., through **active user groups** that can build bodies of knowledge and facilitate informal, peer-based knowledge sharing and problem solving.
DOCUMENTED	ADR UK flagship dataset must have a publicly available user guide and data dictionary that meet agreed minimum standards and are updated over time to reflect improvements. Documentation should also be available within the relevant ADR UK trusted research environment.	i. A **data dictionary and user guide** that meet the ADR UK minimum standards of documentation including: • a default, user-friendly dataset name used on all platforms and references to support promotion, findability and avoid confusion (consistent with naming convention standards) • a statement to confirm the user guide and data dictionary have met and been validated against the minimum standard as well as a commitment to update and improve the documentation over time. • a formal mechanism for researchers to provide feedback on the documentation (e.g., online form) to be taken into account during an annual review and clearly signposted to researchers via the documentation. • the location of the one authoritative version of the documentation, with clear version control, from which all other platforms will link to • the delegated responsible owner who will update and improve the documentation	**Documentation should be updated** annually or with every data refresh whichever is sooner. **Dataset name should be updated** with every data refresh as per the naming convention standards. The documentation owner should regularly review and action feedback from researchers (received through a clear mechanism) about their experience of the data and documentation, seeking to improve the quality of the documentation and underlying data (SEE CURATED). Consideration should be taken to extend the documentation to better articulate the emerging body of knowledge around the data including informal researcher insights (e.g., quirks, known issues), key changes to relevant policy, collection practices that impact the interpretation of the data or to better articulate data quality metrics including linkage quality/accuracy and implications for users of the data. **Note**: to complement the minimum standards for documentation a best practice guidance document

Guiding Principle	Guiding principle description	Minimum standard: evidence for self-assessment	Enduring research asset requirements
		and validate it against the minimum standard • confirmation that documentation is available inside the relevant ADR UK secure environment • a digital object identifier for researchers to cite in their research outputs to support impact tracking ii. A commitment to work towards the enduring research asset requirements over the investment period 2026-2031 including a nominated individual or organisation responsible for coordinating this guiding principle. **Note**: the minimum standards of documentation are due for review by March 2026 and can be revised to capture refinements led by the Flagship Dataset Framework. Where appropriate allow for differences in documentation requirements for pre-linked and linkable datasets.	for documentation could be developed to support more consistent and effective documentation.
ENHANCED	ADR UK flagship datasets are expected to be refreshed by design and enhanced to improve the utility of these enduring research assets over time.	i. A commitment to work towards the enduring research asset requirements over the investment period 2026-2031 including a nominated owner for this guiding principle. The pace of this may vary depending on the maturity of the dataset and data owner relationship but the expectation is that refreshing and enhancing data are an explicit direction of travel. ii. Where available, the provision of any known deliverables to refresh and/or enhance the data or an outline plan that moves towards regular refreshes/enhancement including a	Regularly **refresh** the data at an agreed frequency (e.g. annually) by updating or replacing existing data with the most current data available. Include the frequency and date of the refresh in the documentation, the ADR UK flagship dataset webpages and appropriate catalogue entry. NOTE: the refresh frequency does not need to be annual but should be appropriate for the dataset, the data owner preferences and its corresponding research value. **Enhance** the data over time by improving the quality and usefulness of the existing data by

Guiding Principle	Guiding principle description	Minimum standard: evidence for self-assessment	Enduring research asset requirements
		nominated owner for refreshing and enhancing the data. Milestones will be assessed via routine quarterly highlight reporting.	adding additional linkages or variables. This could include: • aggregate or publicly available data to support the analytical potential of the data such as geographical fields or household units • new derived variables to create research concepts or harmonised variables to simplify data structures and make it easier to use longitudinal data across time and cohorts. It may be more appropriate to manage these through documentation and code to allow the researcher to reproduce the derived variable themselves • new data linkages from other sources to fill data gaps. • simply using an anonymised linkage field that brings multiple data sources together is an act of enhancement as all new data acquisitions have the potential to enhance all other data.
FACILITATED	ADR UK flagship datasets should have associated support, resources and engagement activities to inform, inspire and enthuse researchers throughout the researcher journey to conduct policy relevant and impactful research that addresses key societal challenges.	i. Cleary signposted (e.g., via the documentation) contact information for an individual or organisation with the appropriate dataset expertise and resource to support researchers with their enquiries including answering technical questions about the dataset throughout the project lifecycle and conducting pre-project approval feasibility checking. ii. A commitment to work towards the enduring research asset requirements over the investment period 2026-2031 including a nominated individual or organisation	Inform, inspire and enthuse researchers to use flagship datasets, both individually and collectively, throughout the life cycle of their flagship dataset projects to conduct policy-relevant and impactful research. Coordinate across government and academic partners to deliver solutions for facilitating research. Although solutions for 'facilitating' researcher support are likely to vary by dataset due to differing resources, expertise and ideas - all partners are expected to collaborate across the partnership to share knowledge, iterate and improve practice and drive innovation to support impactful research use of ADR UK flagship data.

Guiding Principle	Guiding principle description	Minimum standard: evidence for self-assessment	Enduring research asset requirements

		responsible for coordinating this guiding principle.
		
		iii. Where available, associated deliverables to be assessed via the routine quarterly highlight reporting process.
		
		Note: this principle is expected to involve significant partner capacity and coordination and collaboration.
			(Knowledge sharing of current practice relevant to facilitating research use of ADR UK flagship data is being collated in the section below.)
			
			Key areas to develop support, resources and engagement activities are as follows:
			
			1. **Researcher support and resources**
			   - Technical support for dataset specific queries
			   - Analytical or research support for guidance on project feasibility and research design including for example requests for aggregated data outputs which describe the viability of potential research questions.
			   - Access to synthetic data (see best practice section below)
			   - Data-set specific training and resources
			   - Code sharing repositories
			   - User groups, discussion forums including data owner led engagement events
			   - Repositories of projects and publications
			   - Data descriptor papers which go beyond metadata in giving more qualitative or evaluative assessments of the data and exemplar analyses or summaries
			   - Publicly available aggregated data that lower the bar to access the data
			
			2. **Stakeholder & public engagement**
			   - Clear mechanism and signposting for researchers to engage with relevant policy and analytical colleagues to shape research ideas and interpret findings.

Guiding Principle	Guiding principle description	Minimum standard: evidence for self-assessment	Enduring research asset requirements
			• Guidance and or training on the range of potential policy impacts and how to deliver on these. • Signposting to an active and relevant public engagement panel for researchers to test and shape research proposals, insights and outputs with those with lived experience.

References

- Stewart, B., Oliver, E., McGrath-Lone, L., (2022). What is 'research-ready' data? A report on the online roundtable event held on 11 February 2022. *What_is_research_ready_data__A_roundtable_report_June_2022.pdf

- Gordon, B., Barrett, J., Fennessy, C., Cake, C., Milward, A., Irwin, C., Jones, M., & Sebire, N. (2021). Development of a data utility framework to support effective health data curation. BMJ health & care informatics, 28(1), e100303. https://doi.org/10.1136/bmjhci-2020-100303

- McGrath-Lone, L., A Jay, M., Blackburn, R., Gordon E., Zylbersztejn, A., Wiljaars, L., Gilbert, R., What makes administrative data "research-ready"? A systematic review and thematic analysis of published literature. Int. J. Popul. Data Sci.; 4. 10.23889/ijpds.v7i1.17182022. https://doi.org/10.23889/ijpds.v7i1.1718