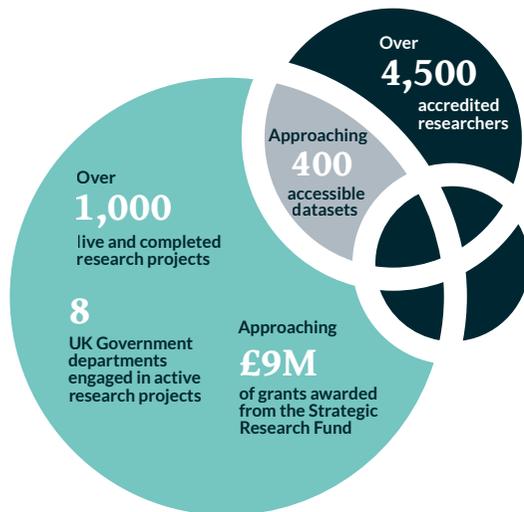# Contents

Introduction

This brochure showcases a collection of recently linked administrative datasets available for research. It is designed to introduce researchers to the wealth of linked public sector data available covering a wide range of topics and policy areas.

What is ADR UK?

ADR UK (Administrative Data Research UK) is a partnership transforming the way researchers access the UK's wealth of public sector data, to enable better informed policy decisions that improve lives.

Over
4,500
accredited researchers

Approaching
400
accessible datasets

Over
1,000
live and completed research projects

8
UK Government departments engaged in active research projects

Approaching
£9M
of grants awarded from the Strategic Research Fund

Administrative data is information created when people interact with public services, such as schools, the NHS, the courts or the benefits system, and is collected by government and other public sector organisations.

Education records

Court records

Health records

Employment & earnings

Census responses
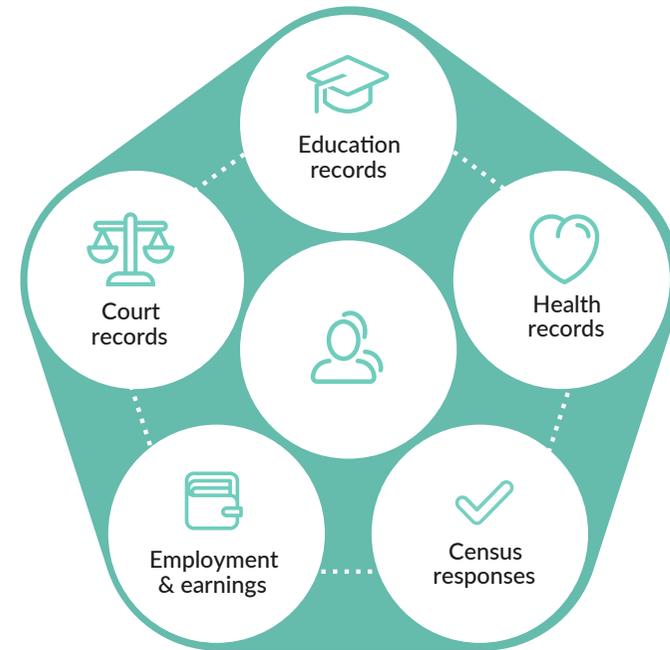
Over the past four years, ADR UK has made significant advances in making de-identified administrative data easier for researchers to access securely, and there is now an impressive collection of new flagship datasets available through our trusted research environments across the UK. In addition, ADR UK is developing further resources including explainer videos, training courses, synthetic data which can be used to develop and test code, and communities of practice in which researchers can share ideas.

The linked administrative datasets featured in this brochure form an exciting first collection of rich, curated data fit for use by researchers across all career levels. ADR UK's commitment to provide extra resources will add to the appeal of these datasets and pave the way for further collections to become available.

Read how ADR UK ensures data is used ethically and responsibly.

Longitudinal Education Outcomes

Summary

The Longitudinal Education Outcomes (LEO) dataset is a world-leading data product created by the Department for Education (DfE). It contains de-identified information on the characteristics, education, employment, benefits, and earnings of members of the British public. This is a unique source of information, with the potential to provide transformative insight and evidence on the longer-term labour market outcomes and educational pathways of (currently) around 38 million individuals. The purpose of the LEO dataset is to enable the UK to gain transformational insights about pathways within and after education, to enhance the life chances of current and future learners through the enhancement of evidence-based decision making.

Overview of the component datasets and linkage

LEO is a relational database comprised of a range of tables and access is restricted on a needs basis.

Timeframe: Jan 1995 – Aug 2019
Update frequency: At least annual
Coverage: England only
Size: ~38 million records

Longitudinal Education Outcomes (LEO) component datasets and linkage

- **National Pupil Database** — School and early years data
- **Individual Learner Record Data** — College and apprenticeship data
- **Higher Education Statistics Agency data** — Higher education data
- **LEO data (HMRC and DWP)** — Earnings, benefits, and tax data
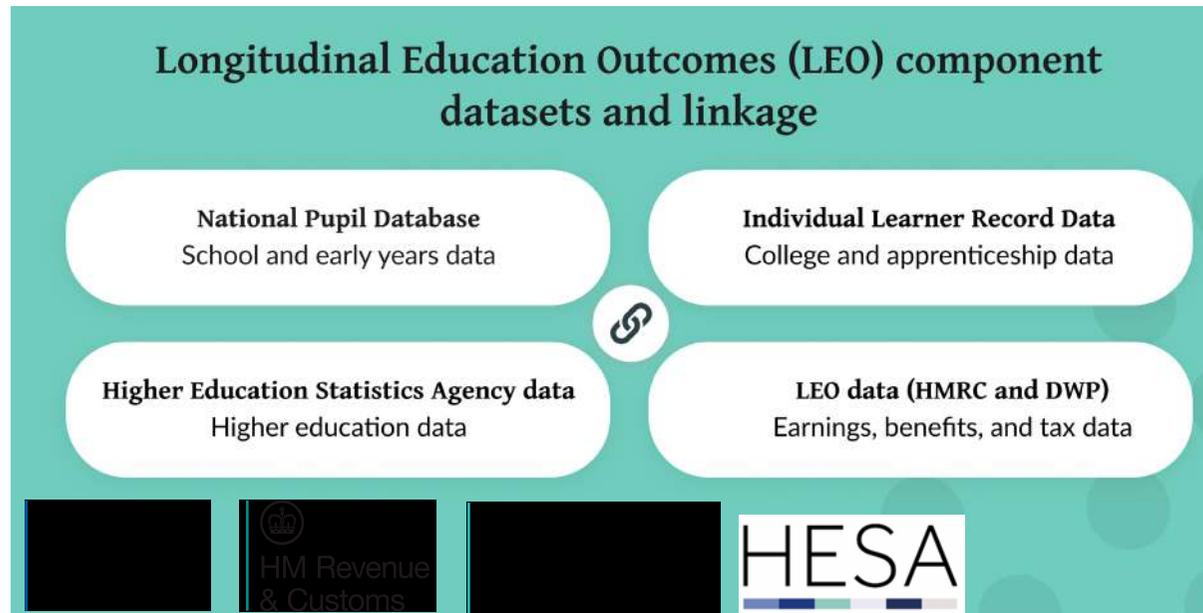
Core documentation

User guide: The LEO user guide is available to [accredited researchers](#) within the [Office for National Statistics (ONS) Secure Research Service](#) and the [LEO Research Community](#).

Data dictionary: Information about variables within the current version of the LEO Standard Extract can be found in the [LEO variable request form](#) (update expected September/October 2022).

Examples of research questions the data could be used to address

- Does attendance at a school with a "Good" Ofsted rating have an impact on a person's labour market outcomes?

- Are pupil referral units (for pupils who have been excluded from school) effective in improving outcomes for young people?

- Does being in the care system lead to poorer labour market outcomes after controlling for other demographic and education factors?

- Which education factors have a positive impact on labour market outcomes of school leavers with poor attainment?

- Does coming from a lower socioeconomic background lead to poorer labour market outcomes after controlling for other demographic and education factors?

- What is the link between location, labour mobility and education and labour market outcomes? At similar education levels, do those who are more mobile have better outcomes?

Examples of research done using the data

The [LEO Research Community](#) library stores a list of published research. Examples from the broad and largely un-tapped field of potential LEO research include:

- Research into [post-16 pathways](#) (journeys of individuals following their state education).

- Research into social mobility and outcomes for students receiving free school meals in England.

Other supporting information and resources

Webinar: The Edge Foundation recently [published a recording](#) showcasing analyses using the LEO dataset.

Guidance & frequently asked questions on how to access the LEO dataset are available from the [ONS Secure Research Service](#).

LEO Library: The LEO Research Community supports data preparation and analysis and hosts a growing library of resources and information. To access the LEO Community you will need a Knowledge Hub account and UK Statistics Authority (UKSA) researcher accreditation.

More information

Find more information about the [LEO programme](#) and the [National Pupil Database](#) from gov.uk. The Higher Education Statistics Agency (HESA) provide information and [user guides for Higher Education data](#) on their website. For specific queries about the LEO programme, please contact [leo.programme@education.gov.uk](#).

Grading and Admissions Data for England

Summary

The Grading and Admissions Data for England (GRADE) is a joint open data initiative by Ofqual, Ofsted, the Department for Education (DfE), and UCAS. Its main objective is to provide accredited researchers with access to data to conduct independent research on the educational, assessment, and admission systems in England, and to facilitate internal analysis.

It will allow researchers to conduct independent high-quality evaluation of the judgements made in awarding grades in 2020, ensuring that lessons are learned. More broadly, the GRADE data will enable research to enhance the quality of the assessment system and produce evidence to inform future education policy, particularly around the fairness of methods for measuring students' attainment, and implications for university admissions processes.
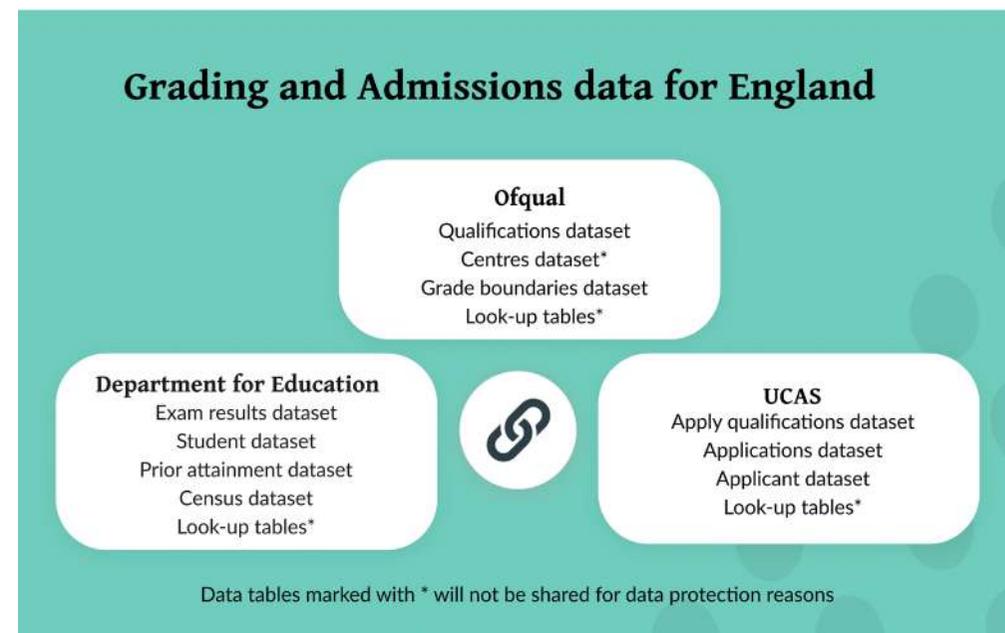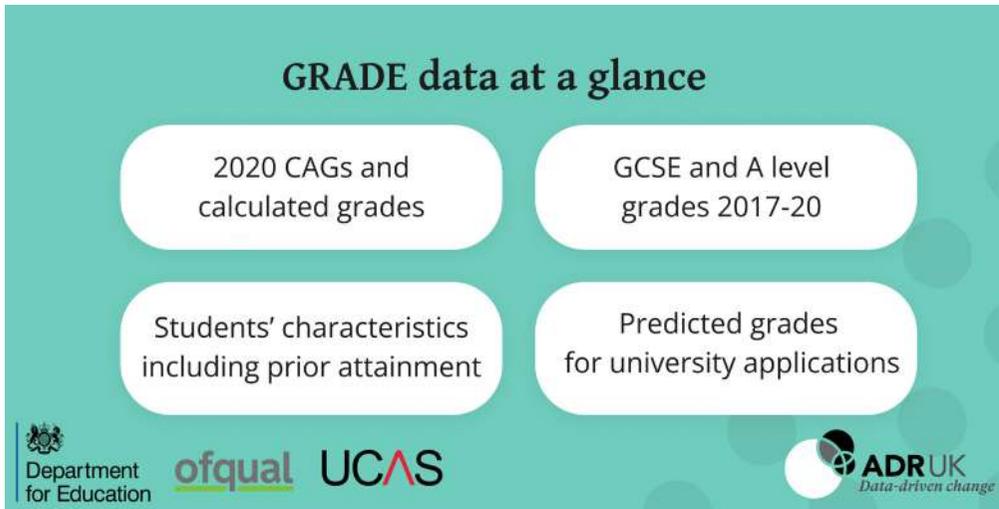
Overview of the component datasets and linkage

Timeframe: Jan-2017 – Jan 2021
Update frequency: Every two years
Coverage: England only
Size: ~ 4 million records

GRADE data at a glance

- 2020 CAGs and calculated grades
- GCSE and A level grades 2017-20
- Students' characteristics including prior attainment
- Predicted grades for university applications

Department for Education · ofqual · UCAS · ADR UK Data-driven change

Core documentation

User guide: A data sharing Framework provides a high-level description of the datasets and sets out how to access the data. A more detailed user guide will be published soon.

Data dictionaries: Data specifications are available from Ofqual, UCAS, and DfE (NPD) on .gov.uk.

Other supporting information and resources

Synthetic data: A low-fidelity synthetic dataset with the same format and structure of the real data, but containing only made-up information based on simulated data, is available by request. Please email data.sharing@ofqual.gov.uk.

Examples of research questions the data could be used to address

- How does the standardisation model introduced in summer 2020 stand up to scrutiny?

- What was the impact of awarding approaches on students and schools and/ or colleges?

- What is the relationship between the 2020 grading data and other measures of attainment (validity related questions)?

Examples of research done using the data

Non-journal publications: Two research reports looking at the awarding of centre assessment grades for GCSEs, AS and A levels in summer 2020.

ADR UK Research Fellowship: ADR UK is funding a Research Fellow to carry out pathfinder research using the GRADE data linkage (call now closed). Keep posted about their upcoming research activities and outputs on the ADR UK website or sign up for updates.

More information

Find more information about the data and its research potential in the ADR UK Fellowship funding call specification (note that this opportunity has now closed but that the information provided about the data is still relevant), the ONS Metadata Catalogue, and the GRADE website.
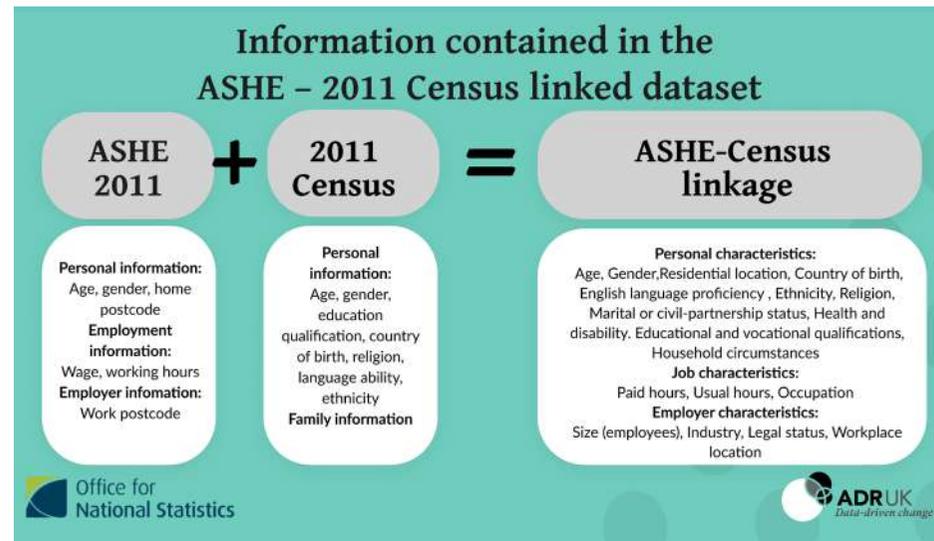
Wage and Employment Dynamics Phase 1

Summary

The [Wage and Employment Dynamics (WED)](#) data linkage project is a collaboration led by researchers from the University of the West of England, aiming to better understand wage inequalities in Britain. Phase one of the project is made up of two parts: 1) Linkage of the Annual Survey of Hours and Earnings (ASHE) data to 2011 Census; and 2) the development of Stata code to enrich the ASHE data. This will allow for insight into the dynamics of wage and employment issues, and how characteristics such as gender, disability, and ethnicity influence these. The ASHE – 2011 Census linkage is expected to be available for access through the ONS Secure Research Service in autumn 2022.

Overview of the component datasets and linkage

1) ASHE – 2011 Census linked dataset

2) Enriched ASHE code

The Wage and Employment Dynamics team has enriched the core ASHE dataset by adding new variables such as minimum wage rates and survey dates and carrying out methodological reviews on the data and sampling. They have also created new weights which account for attrition and selection effects. These are available from 2004. Access to the code will enable research into the dynamics of wage and employment issues, from labour market entry, through job mobility and career progression to retirement. The code created by the WED team will be available within the ONS Secure Research Service on a shared accessible drive. Please note that this code is not part of the ONS ASHE official Statistics output.

Timeframe: 1997 - 2020

Update frequency: Annual

Population: ASHE population spine (attained sample is approximately 0.66% of the entire workforce)

Coverage: England and Wales*

Size: ~175,000 records per year

* Northern Ireland: The Earnings and Employees Study (EES) 2011 for Northern Ireland links together variables from the Annual Survey of Hours and Earnings (ASHE) 2011 with variables from the Census of Population and Housing 2011, and Capital Value data from the Land and Property Services. The dataset takes the form of one pre-linked table.

Core documentation

User guide for ASHE-2011 Census and Code for Enriched ASHE Quick Start Guide: Both these documents can be found on the [Wage and Employment Dynamics website](#).

Data dictionary: Information about the variables will be available from the ONS Secure Research Service Metadata Catalogue.

Other supporting information and resources

Training to use the data: See the [upcoming training opportunities on the data](#), as well as recorded webinars from previous training sessions.

Examples of research questions the data could be used to address

- How do people's earnings change throughout their career? And how does this differ depending on characteristics such as gender, disability, or ethnicity?

- What are the patterns or characteristics of those who do and do not progress out of low paid employment?

- What is the relationship between migration and the labour market?

- What role do employers play in wage inequality?

Examples of research done using the data

Wage and Employment Dynamics [publish research and reports on their website](#).

More information

Find out more about the [Wage and Employment Dynamics](#) project and data. Keep posted about their upcoming events, outputs and activities on the [ADR UK website](#) or [sign up for updates](#). Or [get in touch](#) with the team directly via their website.

> **Tim Butcher, Chief Economist, Low Pay Commission:**
>
> *"This project addresses weaknesses in our evidence base – improving the quality of longitudinal earnings data and extending coverage to a broader range of characteristics – that should enable researchers to give new and innovative insights into the wage and employment dynamics of the lowest paid."*

Data First Datasets and Linkages

Summary

Data First is an ambitious data-linking programme led by the Ministry of Justice (MoJ) and funded by ADR UK. It aims to harness the potential of the wealth of data already created by MoJ, by linking administrative datasets from across the justice system and beyond and enabling researchers within government as well as approved academics to access the data in an ethical and responsible way.

Data First datasets, key features, and linkages

Data First: probation and criminal justice system linked datasets

Magistrates' court defendant data
Timeframe: January 2011-January 2022
Coverage: England and Wales
Size: 15.1 million records

Crown Court defendant data
Timeframe: January 2013-January 2022
Coverage: England and Wales
Size: 1 million records

Criminal courts, prisons and probation linking dataset
This dataset is a lookup table to identify records that refer to the same people across different datasets

Prisoner custodial journey dataset
Timeframe: January 2001-September 2021
Coverage: England and Wales
Size: 1.3 million records

Probation dataset
Timeframe: January 2014-January 2022
Coverage: England and Wales
Size: 1.9 million records

Data First: Family court-Cafcass linked dataset

Family court data
Timeframe: January 2011-
February 2021
Coverage: England and Wales
Size: 16.5 million records

Cafcass data
Timeframe: April 2007-December 2021
Coverage: England (Cafcass)
and Wales (Cafcass Cyrmu)

Data First: Ministry of Justice - Department for Education linked dataset

Police National Computer extract (PNC)
Timeframe: 2000-2017
Coverage: England
Size: 1.65 million records

National Pupil Database (NPD)
Timeframe: Children born after 31
August 1985
Coverage: England
Size: 15 million records

Information about research priorities

The MoJ Areas of Research Interest 2020 sets out the department's critical evidence gaps that could be explored using the Data First datasets.

Data use register

Find out how researchers are using the Data First datasets in the MoJ data requests log, which records all approved external data requests.

More information

More information about the Data First datasets and linkages can be found on pages 14-19 of this brochure, as well as on the ONS Metadata Catalogue, the ADR UK website, and the Data First website. For more specific data, linkage, or metadata queries you can also email datafirst@justice.gov.uk.

> **Lord Justice Fulford, Vice President of the Court of Appeal (Criminal Division) and previous Judicial Lead on Data:**
>
> *"I consider this to be an extremely important project, which has been established according to all of the principles that necessarily apply to work of this kind. In a sentence, it will enable critical research to be undertaken by accredited professionals, whilst protecting the identities of those involved in the individual cases."*
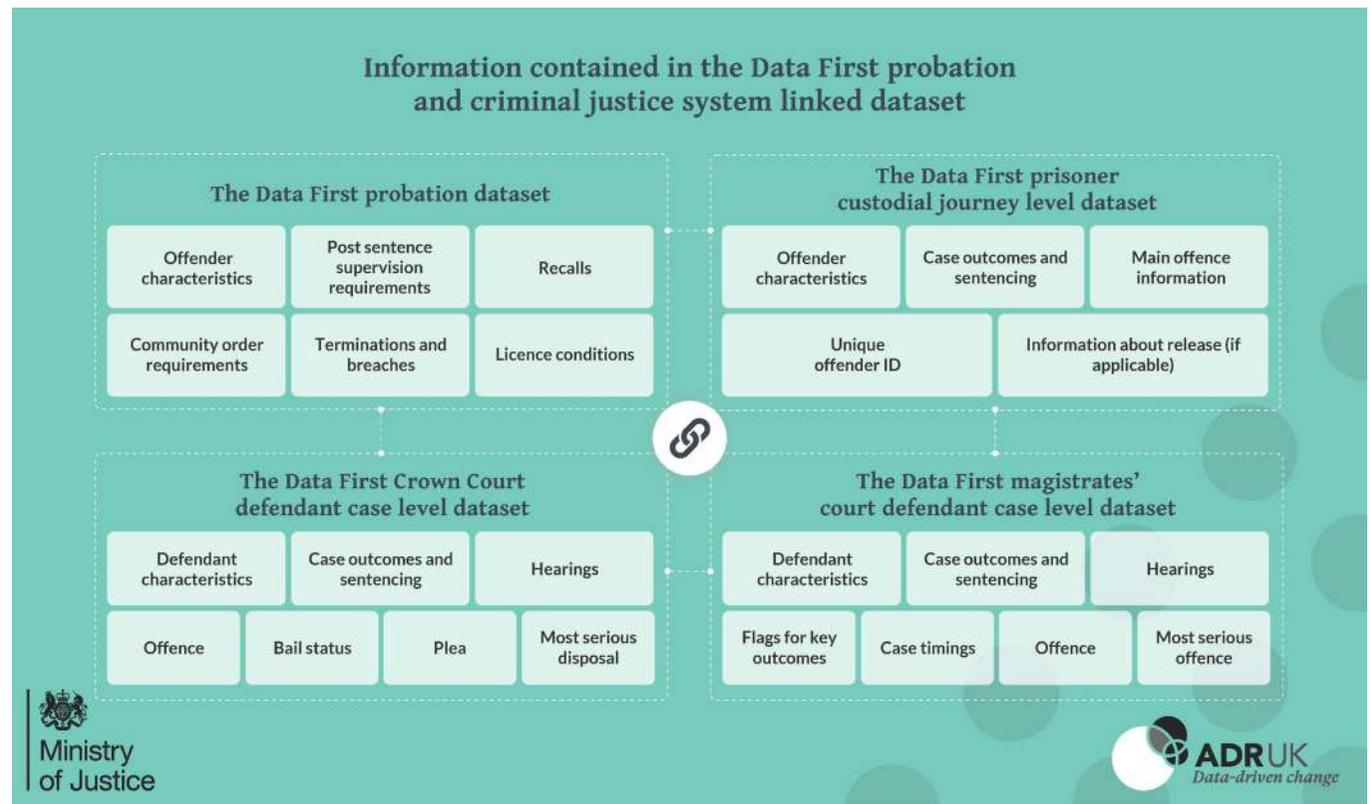
Data First: Criminal Justice System Linked Datasets

Summary

The availability of both magistrates' court and Crown Court data on defendants and cases, as well as data on prisoners' custodial journeys, interactions with probation services, and the ability to link these together, presents an unprecedented opportunity to gain insights on end-to-end user journeys across the entire criminal justice system.

Overview of the component datasets and linkage

Information contained in the Data First probation and criminal justice system linked dataset

The Data First probation dataset

| Offender characteristics | Post sentence supervision requirements | Recalls |

| Community order requirements | Terminations and breaches | Licence conditions |

The Data First prisoner custodial journey level dataset

| Offender characteristics | Case outcomes and sentencing | Main offence information |

| Unique offender ID | Information about release (if applicable) |

The Data First Crown Court defendant case level dataset

| Defendant characteristics | Case outcomes and sentencing | Hearings |

| Offence | Bail status | Plea | Most serious disposal |

The Data First magistrates' court defendant case level dataset

| Defendant characteristics | Case outcomes and sentencing | Hearings |

| Flags for key outcomes | Case timings | Offence | Most serious offence |

Ministry of Justice

ADR UK
Data-driven change

Core documentation

User guide: Data First Introductory User Guide

Data catalogues:

- Magistrates' court defendant case level dataset

- Crown Court defendant case level dataset

- Probation dataset

- Prisoner custodial journey level dataset

- Criminal courts, prisons and probation linking dataset

Other supporting information and resources

Synthetic data: 1000 row, low-fidelity synthetic datasets are available for the magistrates' court and Crown Court datasets. These can be requested by emailing datafirst@justice.gov.uk.

'**Data Explained**': Data Explained are produced by ADR UK Research Fellows discussing the first-hand experiences on the utility and limitations of the data for their research projects, for example by Dr Angela Sorsby. These can be found on the ADR UK project page, and more will be published as they become available.

Information about research priorities

Areas of Research Interest (ARI): The Ministry of Justice ARI 2020 sets out the department's critical evidence gaps.

User Representation Panel report: A report on the Data First User Representation Panel roundtable held 31 January 2022 on the research interest and potential of the criminal justice data.

Examples of research questions the data could be used to address

- Who are the 'repeat users' of the criminal justice system? What are their characteristics? How often do they return? How do outcomes change on each return to the criminal justice system?

- What factors affect the likelihood of different groups receiving different sentences?

- How has the use of non-custodial and custodial sentencing changed over time?

Examples of research done using the data

Non-journal article publications: An exploratory analysis of the magistrates' and Crown Court defendant datasets has been published by the Ministry of Justice.

ADR UK Research Fellows' outputs: Four ADR UK Research Fellows have undertaken pathfinder research using the linked magistrates' and Crown Court defendant datasets. Their publications including blogs and Data Insights can be accessed via the ADR UK project page. Sign up to our newsletter for updates.

More information

More information about the Data First criminal justice datasets and linkages can be found on the ADR UK website and on the MoJ Data First website.

Data First: Family Court – Cafcass Linked Dataset

Summary

The Data First family court - Cafcass linked dataset came about from a data sharing agreement between the Secure Anonymised Information Linkage (SAIL) Databank, Ministry of Justice and Cafcass (Children and Family Court Advisory and Support Service). The de-identified dataset contains information on public and private law cases, adoption cases, marriage and divorce characteristics and their legal outcomes in England and Wales.

Overview of the component datasets and linkage

Information contained in the Data First family court-Cafcass linked dataset

- Unique ID for party
- Divorce, private law and Family Law Act cases
- Public law and adoption cases
- Family law applications and legal outputs
- Marriage and divorce characteristics
- Personal-level information
- Case events, timeliness and number of hearings
- Legal representation

Ministry of Justice   Cafcass   ADR UK Data-driven change

NB: Cafcass, Cafcass Cymru, and the linked Cafcass – Family court data are currently only available to access in the SAIL databank.

Core documentation

User guide: [Data First Introductory User Guide](#)

Data catalogues:

- [Family court](#)
- [Cafcass England](#)

Other supporting resources

Webinar: This [explainer webinar](#) from Professor Karen Broadhurst gives an overview of the Family Justice system and how the family court - Cafcass linked dataset can be used to address key evidence gaps.

Information about research priorities

Areas of Research Interest (ARI): The [Ministry of Justice ARI 2020](#) sets out the department's critical evidence gaps.

Infographic: The Nuffield Family Justice Observatory [children in the family justice system infographic](#) summarises evidence and evidence gaps.

Webinar: Lisa Harker, Nuffield discusses [key evidence gaps in this field](#) in a webinar.

Additional funding opportunities

ADR UK Research Fellowships: At the time of publication, ADR UK is accepting applications to support 15-month research fellowships to carry out pathfinder research using the family court – Cafcass linked dataset. Find out more on the [ADR UK website](#). The deadline to apply is 16.00 on 20 October 2022.

Examples of research questions the data could be used to address

- How do private and public law cases overlap?
- What are the characteristics of repeat users of the family court?
- What is the impact of long-running (i.e., protracted) proceedings on children and families?
- What are the characteristics of cases involving litigants in person (those in court without legal representation)?
- To what extent is there regional variation (by local authority and/or court area) in the way family courts are dealing with cases?
- What types of cases attempt mediation within the family jurisdiction?

More information

More information about this linked dataset can be found on the [ADR UK website](#).

> **Sir Andrew McFarlane, President of the Family Division of the judiciary in England and Wales:**
>
> *"The Data First project is likely to be the most significant step forward in the provision of hard and reliable information about the family justice system for a generation. It breaks entirely new ground and the fruits of the data that will be provided are likely to inform policy and practice and identify important trends for decades to come. Its advent could not be more welcome."*
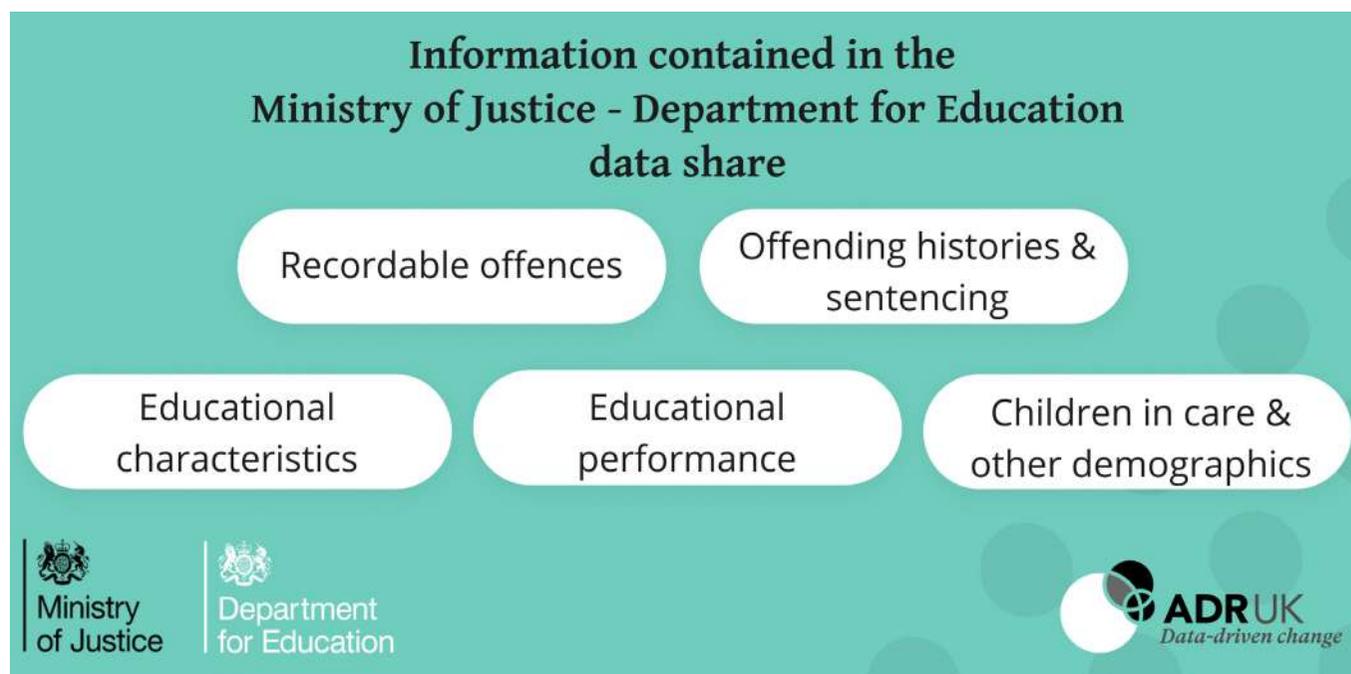
Data First: Ministry of Justice – Department for Education Linked Dataset

Summary

The MoJ-DfE linked dataset allows analysis to understand links between childhood characteristics, educational outcomes and (re)offending. The dataset includes data from Prison, Courts, Police National Computer, National Pupil Database, Looked-After-Children and Children in Need and covers variables such as demographics, offending data, school exclusions, and children in care. Analysis of this data will help develop the evidence base on understanding the relationships between educational and criminal justice outcomes and the drivers of offending. It will assist in identifying the population that requires support through early intervention and evaluating interventions to understand whether they are effective.

Overview of the component datasets and linkage

Information contained in the Ministry of Justice - Department for Education data share

- Recordable offences
- Offending histories & sentencing
- Educational characteristics
- Educational performance
- Children in care & other demographics

Ministry of Justice | Department for Education

ADR UK Data-driven change

Core documentation

User guide: Data First Introductory User Guide

Data catalogues: These can be requested by emailing datalinkingteam@justice.gov.uk and data.sharing@education.gov.uk.

Information about research priorities

Areas of Research Interest (ARI): The Ministry of Justice ARI 2020 sets out the department's critical evidence gaps.

Examples of research questions the data could be used to address

- Is there an association between particular interactions with the education system and offending, and if so, is one of these factors typically the driver?

- How are the relationships between educational and criminal justice outcomes impacted by demographic factors?

- Can education policy have unintended consequences for crime?

Examples of research done using the data

Feasibility study: This study explores the use of the MoJ-DfE linked dataset to evaluate early interventions to prevent violent crime.

Non-journal publications using the data:
- The Ministry of Justice published an examination of the education and social care background of children cautioned or sentenced for an offence.

- The Office for National Statistics published an analysis of the education and social care background of young people who interact with the criminal justice system.

- **ADR UK funded Research Fellows' outputs**: Five ADR UK Research Fellows are undertaking pathfinder research using this dataset. Their publications including blogs and Data Insights will be available via the ADR UK project page. Sign up to our newsletter for updates.

More information

More information about this linked dataset can be found on the ADR UK website and on the MoJ Data First website.

> "
>
> **Dr Matteo Sandi, Research Economist, London School of Economics:**
>
> *"I have done research on education policy and youth crime for a few years now using data from a number of countries, and this is probably the best data source that exists today in the world to study youth crime and links between education and crime."*
>
> "

Growing Up in England Dataset

## Summary

ADR UK, the Office for National Statistics (ONS) and the Department for Education (DfE) are working together to make better use of existing data about children and young people in England. The Growing Up in England (GUIE) dataset will enable a more comprehensive understanding of how factors such as family background and geography shape outcomes. This will allow researchers to shine light on what does and doesn't work in supporting children and young people to thrive and give policymakers valuable evidence for developing services that work better for all. There is intention for a future build to include 2021 Census, a review and update of the education component of the data, and linkage reconstruction.

Overview of the component datasets and Core linkage

Growing Up in England (GUIE) component datasets and linkage
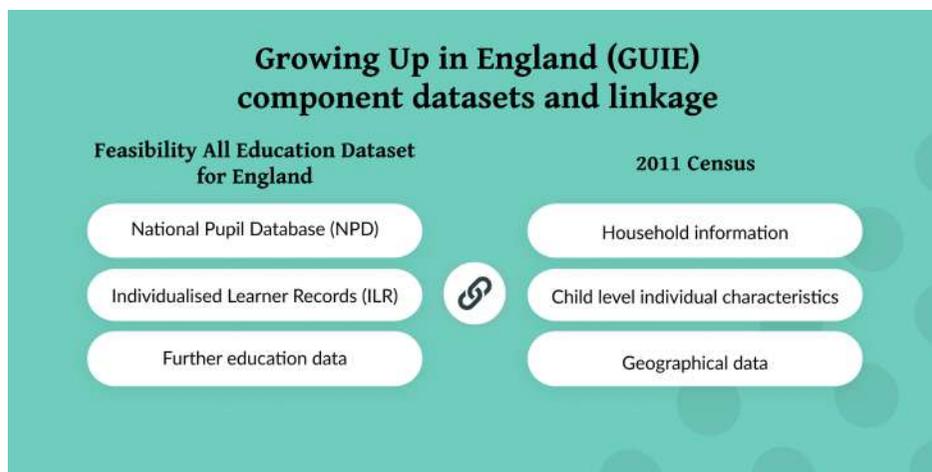
Information contained in the Growing Up in England dataset
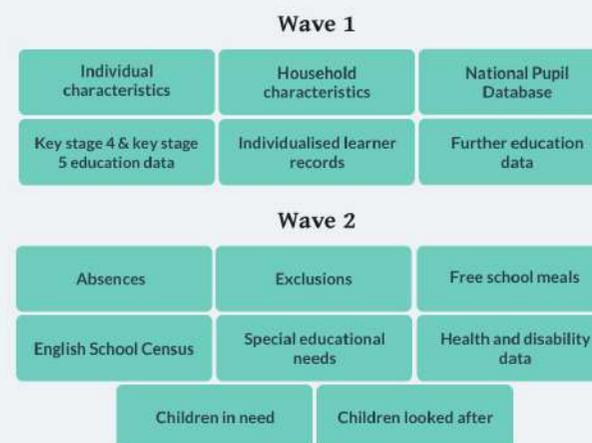
Timeframe:
    Wave 1: Academic years 2001/02 – 2014/15
    Wave 2: Academic years 2010/11 – 2014/15

Update frequency: Static

Population: Key Stage 4 and Key Stage 5 records for children and young people in state-funded schools and colleges (Census 2011-linked dataset includes individuals aged between 14 and 29 years on 31 August 2015).

Coverage: England only

Size: ~7 million records

Core documentation

User guide: A user guide will be made available and will be accessible through the [ONS SRS Metadata Catalogue](). The user guide will also signpost to an accompanying ONS methods publication for Growing Up in England.

Data dictionaries:

- [Wave 1]()
- [Wave 2]() (Children in Need, Children Looked After)
- [Wave 2]() (Exclusions, Absences, and English School Census)

Other supporting resources

AEDE information: read more about the [feasibility All Education Dataset for England (AEDE)]() Note that the AEDE data included in Wave 1 GUIE differs in some respects to what is described in this publication, e.g., HESA data within AEDE is not included in Wave 1 GUIE.
2011 Census information: find out more about the [2011 Census]().

Examples of research questions the data could be used to address

- Is there a relationship between household characteristics and educational attainment?
- What are the underlying drivers of geographical differences in educational attainment?

Examples of research done using the data

Feasibility study: find out more about the data and some proof-of-concept analyses in this [study]() from the ONS. Note that in contrast to the data used in this study, GUIE Wave 1 and 2 only contains individuals who were matched to the AEDE and so does not contain information about parents' characteristics.

Case study: ADR UK have funded a study to explore the [potential of the data and address an evidence gap on educational experiences of Roma, Gypsy, and Traveller children]().

More information

You can find more information about the GUIE dataset on the [ONS Secure Research Service Metadata Catalogue]() and the [GUIE project page]() on the ADR UK website. Read more from the ONS about the potential for administrative data in understanding qualifications in a [blog]() and a [feasibility study]().

Children Looked After Datasets Wales

Summary

The Welsh Government collect several datasets relating to children receiving services from social care, and most of these datasets focus on children who are looked after. These datasets contain rich situational and individual-level data items. The information collected in the Children Looked After (formerly named Looked After Children) Census allows the trajectories of individual children through the care system to be objectively described in a good level of detail. These datasets have recently been shared with the Secure Anonymised Information Linkage (SAIL) Databank. Within this trusted research environment, anonymous linking fields can be used to link individuals across datasets from a wide range of health, education, social care and family justice sources.

Overview of the component datasets and linkage

- **Time covered**: 1 April 2001 – 31 March 2021
- **Update frequency**: Annual
- **Geographical coverage**: Wales
- **Approximate sample size**: 310,550 recorded episodes, reflecting data for 37,775 unique children

In addition to the main Children Looked After dataset, the SAIL Databank holds the following subsets:

- Looked After Children: Adoption (LACA)
- Looked After Children: Care Leaver (LACC)
- Looked After Children: Birthday 19 (LACB)
- Looked After Children: Education Qualifications (LACE)

Core documentation

Children Looked After Census guidance notes
Welsh Government. Stats Wales Statistical Return. 2022

Other supporting resources

- Allnatt, G., Lee, A., Scourfield, J., Elliott, M., Broadhurst, K., Griffiths, L (2022) Data Resource Profile: children looked after administrative records in Wales, International Journal of Population Data Science, 7:1 pp. 1-10.

- A Welsh Social Care Data Users Group has also been established (led by researchers at Swansea, Cardiff, and Lancaster Universities) to support use of Welsh administrative data. This group is affiliated with the Children's Social Care Data User Group (CSCDUG) based in England https://cscdug. co.uk/ and ADR Wales.

Examples of studies using children's social care data

- Allnatt G, Elliot M, Scourfield J, Lee A, Griffiths L. Use of linked administrative children's social care data for research: a scoping review of existing UK studies. British Journal of Social Work, 2022; 00, p.1-22.

- SAIL Databank continues to receive and approve access to these new datasets from researchers across Wales and beyond, such as those based in the **Family Justice Data Partnership** and the **Children's Social Care Research and Developmental Centre**.

Information about policy research priorities

The number of children and young people in Wales who require state care has risen in recent years. Outcomes in various domains relating to education, health and wellbeing for these children are poor. The Welsh Government **Programme for Government** 2021-2026 sets out a number of commitments relating to the care of looked after children.

Information about how to access the data

Follow the application process guidelines available at the **SAIL Databank website**.

Where to find more information

- **Stats Wales**

- Websites for the **SAIL Databank** and **CASCADE** will, in forthcoming months, contain additional information.

"

J Scourfield, Professor of Social Work and former Welsh Government specialist policy adviser:

"The availability of children's social care datasets in SAIL brings great potential for policy-relevant research. There are so many questions that can be answered from population-wide data, including what leads to children being in need, in care and on child protection plans, disaggregated by characteristics such as age and ethnicity; and what the longer-term outcomes are, not only for children looked after but all the other children who receive social services"

"

Scottish Government: Looked After Children Longitudinal Dataset

Overall summary

This dataset brings together data of care placements and the legal basis for children in the care of local authorities across Scotland. This allows understanding of changes in the patterns of care over time and geography, and the relationship with children's characteristics. Information on destinations from the care system are also provided in the data.

The Looked After Children Longitudinal Dataset can be linked to a wide range of other data available in Scotland including data on children attending publicly funded schools. See the ADR Scotland Data Catalogue for further information on available education datasets that can be linked for research purposes.

Core documentation

User guide: Looked after children in Scotland - longitudinal data user guide.

Data dictionary: Looked after Children Data Dictionary.

Overview of the Looked After Children dataset

Geographical coverage:
Scotland

Time period:
2008 to 2019

Approximate sample size:
60,000 children

Other points to note:

70,000 approximate episodes of care

147,000 approximate care placements

195,000 approximate legal reasons

Other supporting resources

Webinar recording: a webinar [recording](#) is available providing an outline of the dataset, background information and examples of research use and linkage.

Researcher handbook: read the Scottish Centre for Administrative Data Research's informative [handbook](#) for researchers.

Examples of research done using the data

There are several research projects and studies ongoing including:

- [Permanently Progressing](#)
- [Investigating Health Outcomes for Care Experienced Children](#)
- [Growing up in kinship care](#)
- [Stability of patterns of care for Children in Scotland](#)

You can also read research reports on:

- [Inequalities in dental health needs and access to dental services among looked after children in Scotland](#)
- [Infants Born into Care](#)

How to access the data from the Scottish Government

Data access requests to the Scottish Government are made via the Statistics Public Benefit and Privacy Panel (Statistics PBPP).

Please read the [Scottish Government guidance](#) on requesting data.

A completed [application form](#) should be submitted to statistics.enquiries@gov.scot.

Data use register of approved projects

A register of approved projects is [published](#) by the Statistics PBPP.

Information about policy research priorities

The Scottish Government published the [Getting it right for looked after children and young people strategy](#) in 2015, which outlines the steps being undertaken to improve the experiences and outcome of all looked after children.

You can find out more about existing evidence and gaps in the evidence base in the [Evidence Framework (independent care review)](#).

More information about the Looked After Children Longitudinal dataset and potential linkages can be found on the [Research Data Scotland website](#), [ADR UK website](#), and [Scottish Centre for Administrative Data Research website](#). You can also email adr.datateam@gov.scot.

> **Craig Kellock, Children and Families Statistics Team Leader, Scottish Government**:
>
> "*This significantly enhances the research potential of statistical data on children with experience of care. By making available anonymised individual-level data, academic research is possible exploring young people's journey through the care system. Crucially, the potential to link to other datasets means that analysis of the experiences and outcomes of care-experienced people can be better understood, and this can inform future policy development.*"

Welsh Environmental Dataset

Summary

This dataset contains household-level exposure to multiple aspects of the built environment. The availability of this data and the ability to link it together and to other secondary data sources presents a unique opportunity to gain further insight on identifying the most modifiable aspects of the built environment to inform local authority and government planners and work towards evaluations of interventions.

Overview of the component datasets and linkage

Household Greenspace Views
- **Time covered**: 2010 - 2018
- **Update frequency**: Irregular
- **Geographical coverage**: Wales
- **Approximate sample size**: 1.4 million homes

Greenspace
- **Time covered**: 2010 - 2019
- **Update frequency**: Irregular
- **Geographical coverage**: Wales
- **Approximate sample size**: 1.4 million homes

Walkability
- **Time covered**: 2011 - 2021
- **Update frequency**: Irregular
- **Geographical coverage**: Wales
- **Approximate sample size**: 1.4 million homes

Pollution
- **Time covered**: 2009-2015
- **Update frequency**: Irregular
- **Geographical coverage**: Cardiff
- **Approximate sample size**: 140,000

Alcohol
- **Time covered**: 2006 - 2011
- **Update frequency**: Irregular
- **Geographical coverage**: Wales
- **Approximate sample size**: 1.4 million homes

Examples of research questions the data could be used to address:

- How do built environment exposures vary socio-demographically?
- Where are the inequalities greatest?

Core documentation

User guide: only accessible within the SAIL gateway.

Data dictionary: only accessible within the SAIL gateway. The data dictionary includes definitions of variables, methodological information, and theoretical basis of environmental variables.

Core documentation above available by request following discussion of requirements with SAIL team.

Other supporting resources

- Support from SAIL team on user training needs (bespoke assessment).
- Attendance at SAIL User Forum.
- Project impact management support from SAIL staff.
- Support and mentoring from a thematic research team in environmental data studies.

Examples of research done using the data

- **Research article** linking air pollution and educational attainment data.
- **Research report** linking alcohol outlets and alcohol related harm to health.
- **Research article** The Green and Blue Spaces (GBS) and mental health in Wales e-cohort.

- **Research article** Impact of air pollution on educational attainment for respiratory health treated students.
- **Research article** A best practice framework to measure spatial variation in alcohol availability.
- **Research article** An international comparative study of active living environments and hospitalization for Wales and Canada.

Data use register of approved projects using the data

The Environmental Health Research Centre of Excellence has a **list of all approved projects**.

Information about policy research priorities

In July 2020, Welsh Government published their **policy focus for building healthy environments**.

Information about how to access the data

- Apply to access the data through the **SAIL Databank**:
  - Complete **MRC Safe Researcher Training**
  - Complete independent **Information Governance Review Panel Application**

Where to find more information

More information can be found on the **SAIL Databank website** and **Environment and Health Research Centre** webpage.

Additional Datasets Expected 2022-2023

The following linked datasets are expected to be ready for accredited researchers to use by the end of 2023, together with additional support materials. Sign up to receive updates as they happen and follow @adr_uk on Twitter.

Dataset name	Description	Expected availability	Further information
AD|ARC (Administrative Data | Agricultural Research Collection)	The AD|ARC data linkage aims to integrate the human dimension with data on farming activities, to better understand the demographic, health, education and economic characteristics of farm households associated with different types and sizes of farm businesses.	TBD	AD|ARC
2011 Census, Benefits and Income	This dataset links 2011 Census data to benefits and income data, including outcomes by race and ethnicity across the public sector.	December 2023	
Public Health Research Data	This dataset includes the Public Health Research Database and is a linked asset which currently includes 2011 Census data; Mortality Data; Hospital Episode Statistics; GP Extraction Service Data for Pandemic Planning and Research data.	Available in catalogue (no user guide available)	PHRD
ECHILD (Education and Child Health Insights from Linked data)	This dataset links records from the National Pupil Database and Hospital Episodes Statistics. Joining this information together will help us understand how children's education affects their health, and how health affects their education.	December 2022* *subject to data owner approval	ECHILD project page
Business Data for Research	This dataset includes the Northern Ireland Annual Business Inquiry.	September 2023	Business statistics | NISRA
Education Outcomes Linkage	Northern Ireland version of the Longitudinal Education Outcomes (England) dataset.	December 2022	

# Accessing Data from the ADR UK Partnership's Trusted Research Environments

Much of the linked administrative data referred to in this brochure uses the **Digital Economy Act 2017** (DEA) as the legal gateway for access. For secure data to be disclosed for public good research purposes under the DEA, a research project must be accredited to confirm that it meets the legal requirements for safe and secure data access. This is assessed by the UK Statistics Authority (UKSA) which accredits research projects for **data accessed under the DEA**.

The diagram below depicts the basic process for accessing data under the DEA. Where a different legal gateway is needed, the process for access also might be different. Please check with the relevant trusted research environment for more detailed information or see the **ADR UK website**.

Applicants need to become an **accredited researcher** in order to access the data. Safe researcher training is provided for researchers who need it, and this can be done through any of the ADR UK partnership's trusted research environments. Researcher accreditation is needed before the research proposal reaches the UKSA accreditation approval stage (see below), so a researcher could submit a research project application to UKSA (via the appropriate trusted research environment) to request access to the data before they become an accredited researcher but will only be granted access following accreditation.

View the administrative datasets available across the ADR UK partnership:

- ONS Metadata Catalogue
- NISRA statistics catalogue
- Research Data Scotland catalogue
- SAIL Databank catalogue
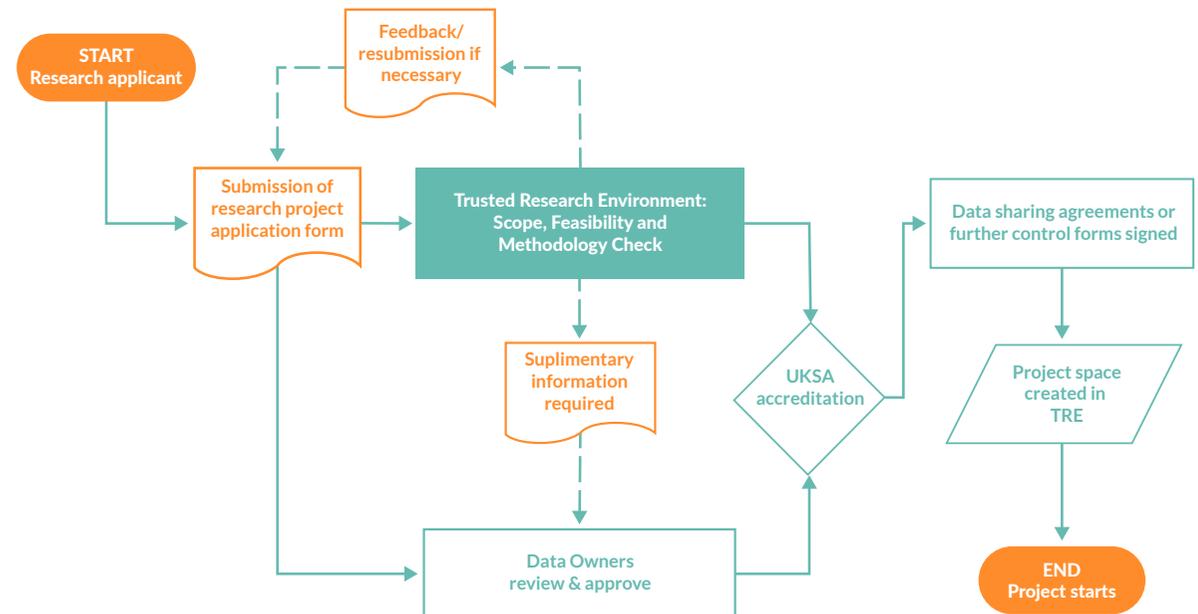
The ADR UK Partnership
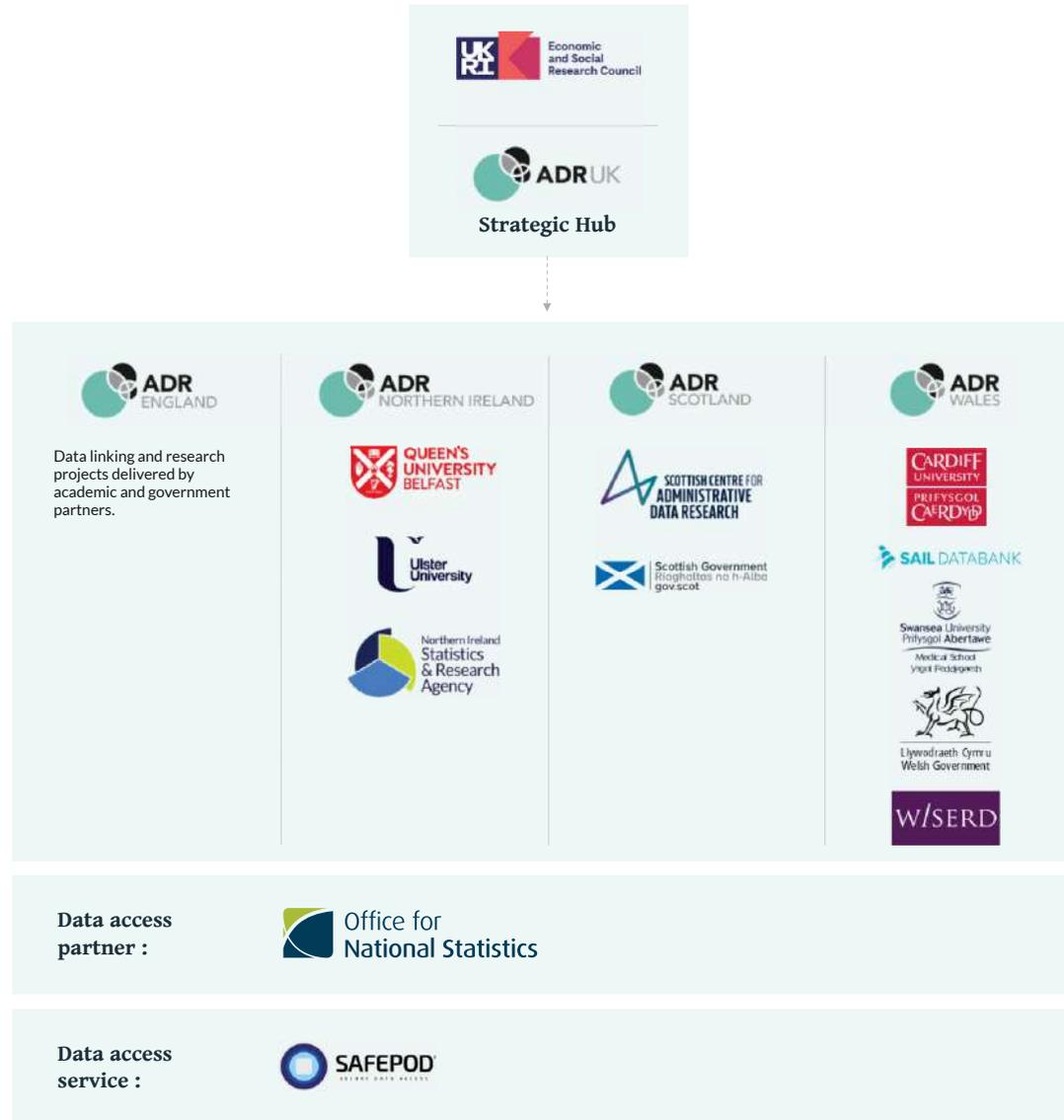
ADR UK (Administrative Data Research UK) is a partnership transforming the way researchers access the UK's wealth of public sector data, to enable better informed policy decisions that improve people's lives.

To find out more, visit adruk.org

Get in touch with us via:

✉ hub@adruk.org
🐦 @adr_uk

ADR UK
Economic and Social Research Council
Polaris House
North Star Avenue
Swindon SN2 1UJ