Data source	Summary	Subject/ participant level	Data access	Data holder
Annual Population Survey (APS)	The APS is an annualised version of the LFS with four quarters of LFS data combined into one annual dataset. Datasets are produced quarterly covering the most recent four quarters. However the datasets are purely cross-sectional and each individual is only represented in one quarter within each APS dataset.	Individuals	Access is free. Users must be registered with the UK Data Service and comply with their conditions of use. There are four types of data access. Standard access: data is anonymised and users are required only to register their project. Special conditions: depositor permission may be required or users may need to agree to a special condition during the download process. Special licence: datasets containing more detailed information may require users to submit special licence applications. Controlled data: datasets containing sensitive information can only be accessed either via the on-site Safe Room or using the online Secure Lab to access the data remotely. The Office for National Statistics: Social Survey Division, who deposited the data, may be informed about usage.	ONS
Annual Survey of Hours and Earnings (ASHE); formerly New Earnings Survey (NES).	ASHE is a longitudinal study tracking pay, hours worked and pensions for a sample of adults aged 16+ who are paid through PAYE in the UK since the 1970s.	Individuals, Institutions/ organisations	Access is free. Users must be registered with the UK Data Service and comply with their conditions of use. There are four types of data access. Standard access: data is anonymised and users are required only to register their project. Special conditions: depositor permission may be required or users may need to agree to a special condition during the download process. Special licence: datasets containing more detailed information may require users to submit special licence applications. Controlled data: datasets containing sensitive information can only be accessed either via the on-site Safe Room or using the online Secure Lab to access the data remotely. The Office for National Statistics, who deposited the data, may be informed about usage. In addition, the Service is required to request explicit permission from the data owner prior to providing the researcher with access to the data.	Office for National Statistics
Aspires/Aspires 2	Aspires and Aspires 2 are a two part research project tracking a cohort of young people in England aged 10-18. It explores attitudes, educational performance and choices around science and science careers. Participants include young people and their parents, and the impact of family and school aspirations and guidance around science and the individual are also explored.	Individuals, Parents	All project data will be made available via the UK Data Service at the end of the project. Access is free. Users must be registered with the UK Data Service and comply with their conditions of use. There are four types of data access. Standard access: data is anonymised and users are required only to register their project. Special conditions: depositor permission may be required or users may need to agree to a special condition during the download process. Special licence: datasets containing more detailed information may require users to submit special licence applications. Controlled data: datasets containing sensitive information can only be accessed either via the on-site Safe Room or using the online Secure Lab to access the data remotely.	Aspires and Aspires 2 were administered and coordinated by the School of Education, Communication & Society at King's College London; it moved to the UCL Institute of Education in March 2017.
Avon Longitudinal Study of Parents and Children (ALSPAC), Children of the 90s; ALSPAC-G2, Children of the Children of the 90s, COCO90s	ALSPAC is a longitudinal study monitoring health and school-based education outcomes for a cohort in Bristol from birth to the present. It also tracks their parents/carers and grandparents to a limited extent. ALSPAC-G2 is a second stage of the research emerging as members of the original young cohort start to have children of their own. These children are now being recruited into the study to be monitored as well to give an intergenerational picture.	Individuals, children, parents, grandparents (ALSPAC-G2), information from staff at Schools attended by youth cohort.	Applications to use the resource are welcome from all 'bona fide' researchers (using the MRC definition) and not for personal or commerical gain. An application form is submitted online and is then considered by the Executive Committee. There is a charge to access the data. Some elements of data are held by the UK Data Service. this access is free. Users must be registered with the UK Data Service and comply with their conditions of use. There are four types of data access. Standard access: data is anonymised and users are required only to register their project. Special conditions: depositor permission may be required or users may need to agree to a special condition during the download process. Special licence: datasets containing more detailed information may require users to submit special licence applications. Controlled data: datasets containing sensitive information can only be accessed either via the on-site Safe Room or using the online Secure Lab to access the data remotely. The University of Britsol who deposited the data, may be informed about usage.	The Principal Investigators of ALSPAC are Professor George Davey Smith and Professor Paul Burton at the University of Bristol.
British Cohort Study 1970, BCS70	The BCS 1970 is a longitudinal study tracking a cohort of young people in England, Scotland and Wales from birth to age 42 and on-going. Participants also include parents, teachers and head teachers. The study explored health, educational and social development, as well as outcomes later in life (into middle age).	Individuals, children, mother.	Access is free. Users must be registered with the UK Data Service and comply with their conditions of use. There are four types of data access. Standard access: data is anonymised and users are required only to register their project. Special conditions: depositor permission may be required or users may need to agree to a special condition during the download process. Special licence: datasets containing more detailed information may require users to submit special licence applications. Controlled data: datasets containing sensitive information can only be accessed either via the on-site Safe Room or using the online Secure Lab to access the data remotely. The University of London. Institute of Education. Centre for Longitudinal Studies, who deposited the data, may be informed about usage.	University of London. Institute of Education. Centre for Longitudinal Studies

Data source	Summary	Subject/ participant level	Data access	Data holder
British Council, Next Generation	Next Generation is a global research series commissioned by the British Council. So far the research is being undertaken in nine countries including the UK. The first UK instance occurred in 2011 and the next UK instance occurred in 2016. It is a snapshot of the attitudes and experiences of undergraduates aged 19-21 years towards international opportunities and global issues.The research comprises desk research, a survey and focus groups. It tracks attitudes over time but not individuals.	Indivdiuals	This data is currently not publically available but the British Council are reviewing whether and how they could make this data available in the future.	British Council.
British Household Panel Survey (BHPS), 1991-2009, Understanding Society, United Kingdom Household Longitudinal Study; (UKHLS), 2009-2015	BHPS/UKHLS (1991 to present) constitute a longitudinal survey to understand social and economic change at the individual and household level in Britain by collecting detail information about employment, income, health, education, activities, political beliefs, wellbeing, environment, social networks, and other personal characteristics.	Individuals, Families/households	Access is free. Users must be registered with the UK Data Service and comply with their conditions of use. There are four types of data access. Standard access: data is anonymised and users are required only to register their project. Special conditions: depositor permission may be required or users may need to agree to a special condition during the download process. Special licence: datasets containing more detailed information may require users to submit special licence applications. Controlled data: datasets containing sensitive information can only be accessed either via the on-site Safe Room or using the online Secure Lab to access the data remotely. The University of Essex. Institute for Social and Economic Research, who deposited the data, may be informed about usage.	University of Essex. Institute for Social and Economic Research
British Social Attitudes Survey (BSA)	BSA is an annual longitudinal survey which started in 1983 and monitors social and political attitudes among adults in Great Britain. The survey comprises a set of core questions collecting demographic information and some key political attitude questions measuring aspects such as agreement with authoritarian/libertarian beliefs. It is not a cohort survey but selects a new sample intended to be representative of the Great Britain population in certain measures. This sample typically comprises 3,000+ individuals.	Individuals	Access is free. Users must be registered with the UK Data Service and comply with their conditions of use. There are four types of data access. Standard access: data is anonymised and users are required only to register their project. Special conditions: depositor permission may be required or users may need to agree to a special condition during the download process. Special licence: datasets containing more detailed information may require users to submit special licence applications. Controlled data: datasets containing sensitive information can only be accessed either via the on-site Safe Room or using the online Secure Lab to access the data remotely. NatCen Social Research, who deposited the data, may be informed about usage. Access to the data requires Athens registration. Data is archived around twelve months after the completion of fieldwork (giving time for analysis and reporting). The data can also be accessed online at BritSocAt www.britsocat.com. BritSocAt provides a free, searchable text database of British Social Attitudes questionnaires and allows users to conduct online analysis of the dataset, either on the entire sample or using a range of socio-demographic break variables.	NatCen Social Research
Business and Community Interaction survey (HE-BCI)	HE-BCI collects survey and quantitative data from Higher Education Providers (HEP) regarding their interactions with businesses and the community. The survey has taken place on an annual basis since 1999. Topics include publically funded research, regeneration, intellectual property, services for business and the community, and public events for the community.	Higher Education Provider (HEP)	Some aggregated data tables are available free on the HESA website. HESA will also provide very minor data queries for free but charge a fee for more substantial requests. The charge depends on the size and complexity of the query but there is a minimum charge of £150 plus VAT. Discounts are available for higher education providers (20%) and for data requests relating to academic research (50%).	HESA
Census: Population census, and Microdata	The UK Census is undertaken every 10 years, with the most recent being on 27 March 2011. Its purpose is to collect population and other statistics to inform planning and resources allocation.	Individuals; Families/households .	Access is free and available from the ONS Census website. Aggregated data is available from the UK Data Service and access is free. This data is available to all and users are not required to be registered with the UK Data Service as it is licensed under the terms of the Open Government Licence. Microdata samples (individual level, and household level) can be accessed the Virtual Microdata Laboratory (VML) facilities at ONS sites in Newport, Titchfield, and London. Researchers must apply to ONS to become an Approved Research to be permitted to access the data. Anonymised versions of the 1991 and 2001 Microdata (referred to as SARS) can be accessed via the UK Data Service and are available for download in accordance with UK Data Service terms and conditions.	ONS, UK Data Service

Data source	Summary	Subject/ participant level	Data access	Data holder
Census: Neighbourhood statistics	The Neighbourhood Statistics Service (NeSS) was established in 2001 by the UK's Office for National Statistics (ONS) and the Neighbourhood Renewal Unit (NRU). Neighbourhood Statistics contains over 300 datasets, covering Health, Housing, Education, Deprivation, Age, Ethnicity and 2001 and 2011 Census data; around 1 billion counts of information.	Data down to neighbourhood level (often to areas of around 250 people in size) for a range of administrative areas including 'Local Authority' and Education Authority'.	Access is free and available from the ONS Neighbourhood Statistics website.	ONS
Census: ONS Longitudinal Study	ONS Longitudinal Study is a complete set of census records for individuals, linked between successive censuses, together with data for various life events such as birth of a child to the mother, widowhood and cancer diagnosis. It relates to a sample of the population of England and Wales. It now includes records for over 950,000 study members.	Individuals; Families/households .	In order to ensure the confidentiality of individuals in the LS data sample, researchers must apply to access the LS. This invovles applying for Researcher Accreditation and then submitting a Research Project Proposal, or an Exploratory Analysis Proposal, as well as an LS supplementary form. Data may only be analysed at a Virtual Microdata Laboratory (VML) which are situated at ONS offices in London, Titchfield and Newport. Alternatively,an aggregated dataset may be specified (which will have a limited number of values) to be released for researchers to analyse at their place of work ie outside of the controlled VML environment.	The Centre for Longitudinal Study Information and User Support (CeLSIUS), UCL.
Citizenship Education Longitudinal Survey (CELS) ; also incorporates Citizens in Transition: Civic Engagement among Young People, CiT;	CELS and CiT constitute a longitudinal study exploring the impact of citizenship education on a cohort of young people in England from 2001 onwards. Participants include young people, teachers and school leaders.	Individuals: young people, teachers, school leaders.	CiT data: Access is free. Users must be registered with the UK Data Service and comply with their conditions of use. There are four types of data access. Standard access: data is anonymised and users are required only to register their project. Special conditions: depositor permission may be required or users may need to agree to a special condition during the download process. Special licence: datasets containing more detailed information may require users to submit special licence applications. Controlled data: datasets containing sensitive information can only be accessed either via the on-site Safe Room or using the online Secure Lab to access the data remotely. NFER, who deposited the data, may be informed about usage.\nCELS data: some parts of the CELS data may be made available in future e.g. parts of the cross-sectional surveys, but different stages of the research have been funded by different organisations, ie DfE and ESRC, and require separate permissions to allow sharing so it may not be possible to share all the datasets.	Since 2013 CELS has been administered by LLAKES Centre for Research on Learning and Life Chances, between 2001 and 2010 it was administered by NFER.\nIn 2011-2012, NFER conducted the Citizens in Transition research which followed up the CELS cohort at age 20. The CiT data is deposited with the UK Data Service but no CELS data has been deposited with them.
Effective Pre-School, Primary and Secondary Education project (EPPSE)	EPPSE is a longitudinal study tracking a cohort of young people in England from age 3 in 1997 to six months after year 11 in 2013 (age 16). Participants include parents and teachers of the young person. The study explores influences from pre-school to secondary school including parents, environment and schooling upon the post-16 destinations of the young cohort.	Individuals, Families/households	Access is free. Users must be registered with the UK Data Service and comply with their conditions of use. There are four types of data access. Standard access: data is anonymised and users are required only to register their project. Special conditions: depositor permission may be required or users may need to agree to a special condition during the download process. Special licence: datasets containing more detailed information may require users to submit special licence applications. Controlled data: datasets containing sensitive information can only be accessed either via the on-site Safe Room or using the online Secure Lab to access the data remotely.\nThe University of London. Institute of Education. Centre for Longitudinal Studies, who deposited the data, may be informed about usage.	University of London. Institute of Education
Employer Perspectives Survey	The Employer Perspectives Survey is a survey involving 18,000 employers, across private, public and non-profit sectors throughout the UK. Themes include work experience; apprenticeships; employer skills sourcing; employer interaction with training providers; recruitment of young people and leavers. The survey started in 2007 and has recurred biannually from 2010 to present.	Employer	Access is free. Users must be registered with the UK Data Service and comply with their conditions of use. There are four types of data access. Standard access: data is anonymised and users are required only to register their project. Special conditions: depositor permission may be required or users may need to agree to a special condition during the download process. Special licence: datasets containing more detailed information may require users to submit special licence applications. Controlled data: datasets containing sensitive information can only be accessed either via the on-site Safe Room or using the online Secure Lab to access the data remotely.\nUK Commission for Employment and Skills (UKCES) or equivalent, who deposited the data, may be informed about usage.	UK Commission for Employment and Skills (UKCES)

Data source	Summary	Subject/ participant level	Data access	Data holder
ESPON Database	The ESPON 2013 Database provides fundamental regional information provided by ESPON projects and EUROSTAT. It provides access to the following data categories: regional, local, urban, neighborhood (candidate countries), world, grid and historical data. Most of the datasets and information produced are publically available and freely accessible. The data and indicators cover the European Union, Switzerland, Norway, Iceland and Liechtenstein (ESPON space). The regional and local data and indicators use the Nomenclature of Territorial Units for Statistics (NUTS) references.	Country level as well as EU NUTS (Classification of Territorial Units for Statistics) regions. For examples, NUTS 1 level areas for England correspond to the former Government regions.	Most of the datasets and information produced are publically available and freely accessible. Users may browse and download data anonymously, but users who register with the portal are permitted access to more datasets.	ESPON holds the data, although data derived from sources which aren't ESPON projects are also available with their original depositor such as EUROSTAT.
European Longitudinal Study of Parents and Children (ELSPAC)	ELSPAC is a longitudinal study monitoring health and school-based education outcomes through participating research groups in ten countries. It also tracks their parents/carers and grandparents to a limited extent.	Individuals, children, parents, grandparents.	Separate arrangments for each country. The Isle of Man ELSPAC data can be accessed through the University of Bristol who hold the UK study data (ALSPAC).	There is no central ELSPAC dataset. Ownership for each study is held separately within each country. In 2014, the University of Bristol took on ownership of the Isle of Man ELSPAC data which will be managed by the Isle of Man Data Governance Committee.
eurostat	Eurostat is the statistical office of the European Union which collects data and publishes statistics at a European level to enable comparisons between countries and regions. Data is collated from national administrative data and surveys. Eurostat's role is to consolidate the data and ensure they are comparable, using harmonized methodology according to the European statistics Code of Practice. This data is then made available via an online database. Statistics are available for a wide range of themes including economy and finance, population and social conditions; industry trade and services; and several other areas.	Country	Free access through the Eurostat website database. Access to Microdata which includes data at an individual, household or employer level can be accessed for scientific purposes. To access this data, researchers must first apply to have theirresearch organisation recognised as a research entity and then apply for access to the data. Anonymised data files will be provided for researchers or non-anonymised data can be accessed through Eurostat's 'safe centre'.	Eurostat.
Families and Children Study (FACS); formerly known as the Survey of Low Income Families (SOLIF)	FACS is a longitudinal study tracking families over a eight to ten year period (1999 to 2009) with a focus on income and/or benefits received and health and educational outcomes for children and use of local services. Children aged 10 - 15 participate in later waves of the survey.	Individuals, Families/households, Parents, children aged 11-15.	Access is free. Users must be registered with the UK Data Service and comply with their conditions of use. There are four types of data access. Standard access: data is anonymised and users are required only to register their project. Special conditions: depositor permission may be required or users may need to agree to a special condition during the download process. Special licence: datasets containing more detailed information may require users to submit special licence applications. Controlled data: datasets containing sensitive information can only be accessed either via the on-site Safe Room or using the online Secure Lab to access the data remotely. The Department for Work and Pensions, who deposited the data, may be informed about usage.	Department for Work and Pensions

Data source	Summary	Subject/ participant level	Data access	Data holder
Futuretrack Study	Futuretrack was a longitudinal study which aimed to elucidate the link between HE, career decision making and labour market opportunities. It was jointly commissioned by HECSU and UCAS. This data was collected by tracking the 2005/2006 cohort of HE applicants, who were then surveyed four times in the six years which followed. It consists of both quantitative and qualitative data.	Individuals	Access is free. Users must be registered with the UK Data Service and comply with their conditions of use. There are four types of data access. Standard access: data is anonymised and users are required only to register their project. Special conditions: depositor permission may be required or users may need to agree to a special condition during the download process. Special licence: datasets containing more detailed information may require users to submit special licence applications. Controlled data: datasets containing sensitive information can only be accessed either via the on-site Safe Room or using the online Secure Lab to access the data remotely. The University of Warwick. Institute for Employment Research, who deposited the data, may be informed about usage.	Higher Education Careers Services Unit (HECSU)
HESA Student Record	HESA student record data has been collected annually since 1994/1995. This student record data records individual level detail regarding demographic data and educational pathways.	Individual students	Some aggregated data tables are available free on the HESA website. HESA will also provide very minor data queries for free but charge a fee for more substantial requests. The charge depends on the size and complexity of the query but there is a minimum charge of £150 plus VAT. Discounts are available for higher education providers (20%) and for data requests relating to academic research (50%).	Higher Education Statistics Agency (HESA)
HESA Destinations of Leavers from Higher Education (DLHE) survey, initial or early DLHE	The HESA Destinations of Leavers from Higher Education (DLHE) survey has been undertaken annually since 1994/1995. It was previously known as the 'First Destinations Supplement' before 2002/2003. The DLHE record is collected from successful leavers six months after they have finished their studies and provides first phase information about patterns of employment and further study. It includes leavers from publicly funded Higher Education Institutions (plus the University of Buckingham) in the UK. Over time the population covered has expanded and now includes UK, EU and international students.	Individual graduates (data is also collected on the employer if the individual is employed)	Some aggregated data tables are available free on the HESA website. HESA will also provide very minor data queries for free but charge a fee for more substantial requests (that require input from HESA analysts). The charge depends on the size and complexity of the query but there is a minimum charge of £150 plus VAT. Discounts are available for higher education providers (20%) and for data requests relating to academic research (50%). Statutory customer of HESA such as DfE receive the data free, and institutions receive data on their own students. HEIDI: Summary DLHE data (from 2011/12) can also be accessed through HEIDI (Higher Education Information Database for Institutions) service to subscribing HEIs and for not for profit national bodies for non-commercial services who can also subscribe to HEIDI. This enables users to explore data about individual providers and compare of DLHE data (and other HESA and non-HESA collated data such as student record, staff data, finance data, estates,National Student Survey,UCAS data, performance indicators) data by institutions. HEIDI will be decommissioned at the end of April 2017 and will be replaced by HEIDI Plus (which was launched in 2015 which has a new user interface tool, contains more visual and interactive representations of HE related data including DLHE,and contains both current and historical data.	Higher Education Statistics Agency (HESA).

Data source	Summary	Subject/ participant level	Data access	Data holder
HESA Destinations of Leavers from Higher Education (DLHE) Longitudinal survey	The HESA Destinations of Leavers from Higher Education (DLHE) Longitudinal dataset is the second part of the DLHE survey. It explores the destinations of leavers up to three and a half years after they have left their HE provider. The two stage approach therefore provides a longitudinal insight into employment and study patterns, and captures details of activities on a census date (usually November) and also satisfaction with career to date. It is a sample survey of those responding to the initial or early DLHE survey. Data has been collected from graduating cohorts every two years (rather than for every graduating cohort). The first cohort surveyed graduated in 2002/03, and data are currently being collected from the 2012/13 cohort (and the most recent data available are for those who graduated in 2010/11).	Individual graduates (data is also collected on the employer if the individual is employed)	Some aggregated data tables are available free on the HESA website. HESA will also provide very minor data queries for free but charge a fee for more substantial requests. The charge depends on the size and complexity of the query but there is a minimum charge of £150 plus VAT. Discounts are available for higher education providers (20%) and for data requests relating to academic research (50%). Statutory customer of HESA such as DfE receive the data free, and institutions receive data on their own students. Summary L DLHE data are available through the HEIDI (Higher Education Information Database for Institutions) service to subscribing HEIs and for not for profit national bodies for non-commercial services who can also subscribe to HEIDI. HEIDI will be decommissioned at the end of April 2017 and will be replaced by HEIDI Plus (which was launched in 2015) which has a new user interface tool, contains more visual and interactive representations of HE related data including LDLHE,and contains both current and historical data.	Higher Education Statistics Agency (HESA)
Individualised Learner Record, ILR	The Individualised Learner Record (ILR) collects information on learners, education programmes and outcomes/destinations in the Further Education system in England since 2003/4.	Individual	Data is held by the UK Data Service but applications to access the data are submitted through the Administrative Data Research Network (ADRN). To access data through ADRN researchers develop a project proposal in discussion with ADRN and make separate researcher applications establishing their background and expertise. These are reviewed by the Approvals Panel and if the research is approved ADRN will then negotiate access to the data with government departments or other organisations. If the data is released, the researcher participates in ADRN training and can then access the data on-site at one of the ADRN secure data labs.	Since 2013, the ILR has been directly collected and administrated by the Skills Funding Agency. Previously the ILR was owned and governed by the Information Authority (IA), while The Data Service was responsible for the collection of the dataset on behalf of the IA
Labour Force Survey (LFS) Longitudinal Datasets	LFS Longitudinal datasets add a longitudinal dimension to the largest regular labour market survey conducted by ONS. Respondents in the LFS are interviewed five times at quarterly intervals, and the longitudinal datasets link responses for the same individuals between the different quarters, to allow an investigation of changes over time for individuals. There are two datasets – a 2Q dataset which links responses between two successive quarters, and the 5Q dataset which links responses across all 5 quarters.	Individuals	Access is free. Users must be registered with the UK Data Service and comply with their conditions of use. There are four types of data access. Standard access: data is anonymised and users are required only to register their project. Special conditions: depositor permission may be required or users may need to agree to a special condition during the download process. Special licence: datasets containing more detailed information may require users to submit special licence applications. Controlled data: datasets containing sensitive information can only be accessed either via the on-site Safe Room or using the online Secure Lab to access the data remotely. The Office of Population Censuses and Surveys. Social Survey Division, who deposited the data, may be informed about usage.	ONS

Data source	Summary	Subject/ participant level	Data access	Data holder
Life Opportunities Survey (LOS)	The LOS was a longitudinal household survey tracking adults aged 16 and over to track transitions between education, into and out of work, barriers experienced, accessability issues, and experience of discrimination. There were three waves between 2009 and 2014. Where there were children aged 11 to 15 in the household then some questions were asked about them as well. Respondents were also asked for consent to link their data with DWP information.	Individuals	Access is free. Users must be registered with the UK Data Service and comply with their conditions of use. There are four types of data access. Standard access: data is anonymised and users are required only to register their project. Special conditions: depositor permission may be required or users may need to agree to a special condition during the download process. Special licence: datasets containing more detailed information may require users to submit special licence applications. Controlled data: datasets containing sensitive information can only be accessed either via the on-site Safe Room or using the online Secure Lab to access the data remotely. The Office for National Statistics. Social Survey Division, who deposited the data, may be informed about usage. There are two version of each LOS study:standard End User Licence (EUL) agreement, and Special Licence (SL) version. The SL version contains geographic variables (Country, Government Office Region and Local Authority), deciles, age bands, a rural/urban indicator, and more information on household relationships, country of previous residence, medical conditions and occupations. The EUL version contains less detailed variables. For example, there is no spatial unit and age is top-coded at 80. Users are advised to first obtain the standard EUL version of the data to see if they are sufficient for their research requirements. The SL data have more restrictive access conditions than those made available under the standard EUL. Prospective users of the SL version will need to complete an extra application form and demonstrate to the data owners exactly why they need access to the additional variables in order to get permission to use that version	Office for National Statistics. Social Survey Division
Local authority interactive tool (LAIT)	LAIT is an Interactive Excel spreadsheet tool containing linked administrative data regarding services and outcomes for young people up to 19 years in England by local authority and region. This includes information regarding health, education, children's services and destinations at 19. Data sources include: Local Authorities (LAs), DfE, OFSTED, and Department of Health. It can be used to compare numbers and some basic percentages for a selected local authority and other local authorities (LAs), at a regional level, and statistically similar LAs.	Local Authority, Regional Level	Available for free download on the GOV.UK website: https://www.gov.uk/government/publications/local-authority-interactive-tool-lait	DfE
Longitudinal Education Outcomes (LEO)	LEO was launched in 2015. It is an experimental dataset using matched administrative data from different government departments (DfE, HMRC and DWP) to follow individuals through their education and into the labour market.The dataset is still in development and Government are engaging with the sector to understand the data and how best to use it. By late 2017 Government plans to have established a regular cycle of publications drawn from the LEO data. LEO comprises experimental datasets exploring a) employment and earnings of graduates (with data matched for those completing their higher education in 1995 or after); and b) employment and education outcomes of young people after Key Stage 5. To date published work using LEO has tracked a cohort of UK domiciled first degree higher education students who graduated from English HEIs in 2003/04 (and graduates from each year afterwards). It has also tracked a cohort of 16-18 year olds who were entered for A levels (or equivalent level 3 qualifications) during the 2012/13 year and were reported in school and college performance tables.	Individual (graduate or Key Stage 5 student)	The uses of LEO (along with other such linked administrative sources) are strictly governed but one of its uses is to undertake or support research projects aimed at furthering the assessment of policy and effectiveness of education and training by appropriate third parties. Access is therefor provided under strict contractual conditions for research and analysis purposes but the data are anonymised. The guidance stipulates that decisions about access to third parties are subject to a robust approval process by an independent panel; and based on a detailed assessment of who request the data, purpose for which it is requested, and level and sensitivity of data requested. Currently, data is only being released externally to organisations undertaking research under contract to the Department for Education. The Department will review its position on making the LEO data available more widely at the end of 2017.	Department for Education

Data source	Summary	Subject/ participant level	Data access	Data holder
First Longitudinal Study of Young People in England (LSYPE1); Next Steps.	LYSPE is an annual panel survey of young people in England. There are seven waves starting in 2004 when the cohort is aged 13-14, ending in 2010 when the cohort is aged 19-20. The survey monitors personal and family characteristics, attitudes, education and employment.	Individuals; Families/households .	Access is free. Users must be registered with the UK Data Service and comply with their conditions of use. There are four types of data access at UKDS. Standard access: data is anonymised and users are required only to register their project. Special conditions: depositor permission may be required or users may need to agree to a special condition during the download process. Special licence: datasets containing more detailed information may require users to submit special licence applications. Controlled data: datasets containing sensitive information can only be accessed either via the on-site Safe Room or using the online Secure Lab to access the data remotely. The Centre for Longitudinal Studies, Institute of Education, University of London, who deposited the data, may be informed about usage.	Commissioned by the Department for Children, Schools and Families (DCSF).
Second Longitudinal Study of Young People in England (LSYPE2); Our Future.	LYSPE2 is an on-going annual panel survey of young people in England. There will be seven waves starting in 2013 when the cohort is aged 13-14. So far Wave 1 has been completed. The survey monitors personal and family characteristics, attitudes, education and employment.	Individuals; Families/households .	Access is free. Users must be registered with the UK Data Service and comply with their conditions of use. There are four types of data access. Standard access: data is anonymised and users are required only to register their project. Special conditions: depositor permission may be required or users may need to agree to a special condition during the download process. Special licence: datasets containing more detailed information may require users to submit special licence applications. Controlled data: datasets containing sensitive information can only be accessed either via the on-site Safe Room or using the online Secure Lab to access the data remotely. TNS BMRB, who deposited the data, may be informed about usage.	Commissioned by the Department for Education (DfE).
Millennium Cohort Study (MCS); Child of the New Century	The MCS is an on-going cohort study tracking individuals and their family circumstances from birth (2000) to age 17 (2018). The most recent data avilable is wave MCS5 (age 11). The study explores family circumstances, health, education, childcare activities, attitiudes & behaviour, and a couple of questions on future plans/expectations.	Individuals, Families/households	Access is free. Users must be registered with the UK Data Service and comply with their conditions of use. There are four types of data access. Standard access: data is anonymised and users are required only to register their project. Special conditions: depositor permission may be required or users may need to agree to a special condition during the download process. Special licence: datasets containing more detailed information may require users to submit special licence applications. Controlled data: datasets containing sensitive information can only be accessed either via the on-site Safe Room or using the online Secure Lab to access the data remotely. The University of London. Institute of Education. Centre for Longitudinal Studies, who deposited the data, may be informed about usage.	University of London. Institute of Education. Centre for Longitudinal Studies
National Pupil Database (NPD)	The National Pupil Database (NPD) contains detailed information regarding all pupils in schools and colleges in England. It mainly covers the state maintained sector but has some limited data on non-maintained provision and independent schools It covers education and attainment at different phases, pupil characteristics, and school information. It draws data from a number of sources including the Pupil Level Annual School Census (PLASC) covering pupils aged 2 to 19 plus,and Key Stages 1 to 5 results, and data has been collected since 2002.	Individuals; schools	Access is free (although a small charge may be incurred when linking NPD to HESA data, currently £420). Users must complete a request form for an extract from the database and submit this to the NPD data request team. Data requests are categorised into 4 tiers by type of data requested. Tier 1 requests covering the most sensitive and personal data including data that will enable an individual to be identified for matching purposes. Whereas Tier 4 requests contain data such as pupil level attainment. The permisson process can take several months, and Tier 1 requests are assessed by a panel and so can take longer to process. Those making Tier 1 or 2 requests need to have a current Disclosure Scotland basic disclosure certificate.	The Department for Education are responsible for the collection and management of NPD.

Data source	Summary	Subject/ participant level	Data access	Data holder
OECD	OECD data collects national level data submitted from national surveys and administrative data to create datasets on a range of topics where results can be compared between countries. These datasets include education and employment. A number of key indicators have been selected, typically at an aggregated level so that data can easily and reliably be compared across countries, e.g. ISCED catgories for education level.	Country	Free download from the OECD library website.	OECD
Opinions and Lifestyle Survey (OPN), formerly Office for National Statistics (ONS) Opinions Survey or ONS Omnibus Survey. It became the Opinions and Lifestyle Survey in April 2012 following the merger of some elements from the General Lifestyle Survey (GLF), which closed in January 2012.	OPN is a frequent, short survey (monthly/8 waves per year) exploring attitudes and experiences among adults aged 16+ in Great Britain (except Isles of Scilly & Scottish highlands & islands). Some core demographic information is collected but topics change with each survey wave. A random sample is selected for each wave.	household level, family level and individual level	Access is free. Users must be registered with the UK Data Service and comply with their conditions of use. There are four types of data access. Standard access: data is anonymised and users are required only to register their project. Special conditions: depositor permission may be required or users may need to agree to a special condition during the download process. Special licence: datasets containing more detailed information may require users to submit special licence applications. Controlled data: datasets containing sensitive information can only be accessed either via the on-site Safe Room or using the online Secure Lab to access the data remotely. The Office for National Statistics. Social Survey Division, who deposited the data, may be informed about usage.	Office for National Statistics. Social Survey Division
School Performance Tables	The School Performance Tables are a set of aggregated tables showing the attainment and progress at Key Stage 2, Key Stage 4 and for those aged between 16 and 18. Tables can be shown for individual school/colleges, local authorities or for England as a whole.	School/college, local authority level.	Access is free, and via the gov.uk website: https://www.gov.uk/school-performance-tables	Department for Education
Universities and College Admissions Service (UCAS)	UCAS collects data from each annual cycle of university and college applicants. The data contains mostly demographic and application information.	Individual	It is free to access aggregated demographic and application related data tables, which can be downloaded on their website. There is a fee for bespoke data requests. UCAS has recently released nine years of university and college admissions data avaliable via the Administrative Data Research Network (ADRN). To access data through ADRN researchers develop a project proposal in discussion with ADRN and make separate researcher applications establishing their background and expertise. These are reviewed by the Approvals Panel and if the research is approved ADRN will then negotiate access to the data with government departments or other organisations. If the data is released, the researcher participates in ADRN training and can then access the data on-site at one of the ADRN secure data labs.	UCAS

Data source	Summary	Subject/ participant level	Data access	Data holder
UCAS Strobe	The UCAS Strobe service is a data linking service provided by UCAS whereby organisations with details of the population of interest can provide lists of individuals to UCAS and UCAS will provide aggregated information about the numbers applying to university, offers received, and places taken up. The service costs a fee per individual, or a fixed fee for very large orders.	Individuals	It is a paid-for service offered by UCAS	UCAS
Young People's Social Attitudes (YPSA)	YPSA is a subsection of the British Social Attitudes (BSA) survey which was carried out in 1994, 1998 and 2003. Young people aged 12 to 19 years were asked about social and political attitudes, including attitudes to work and future plans/aspirations, in order to compare their views with adult BSA respondents. Some basic demographic/personal characteristics data is also collected. Sample size was typically 500 or above.	Individuals	Access is free. Users must be registered with the UK Data Service and comply with their conditions of use. There are four types of data access. Standard access: data is anonymised and users are required only to register their project. Special conditions: depositor permission may be required or users may need to agree to a special condition during the download process. Special licence: datasets containing more detailed information may require users to submit special licence applications. Controlled data: datasets containing sensitive information can only be accessed either via the on-site Safe Room or using the online Secure Lab to access the data remotely. NatCen Social Research, who deposited the data, may be informed about usage. Data is archived around twelve months after the completion of fieldwork (giving time for analysis and reporting).	NatCen Social Research
Youth Cohort Study (YCS)	YCS is a series of cohort studies following individuals for four years following year 11 with annual interviews/surveys. The series starts in 1985 and the most recent cohort (thirteen) took place 2007-10. The research explores personal and family characteristics, education, employment, attitudes and careers advice	Individuals	Access is free. Users must be registered with the UK Data Service and comply with their conditions of use. There are four types of data access. Standard access: data is anonymised and users are required only to register their project. Special conditions: depositor permission may be required or users may need to agree to a special condition during the download process. Special licence: datasets containing more detailed information may require users to submit special licence applications. Controlled data: datasets containing sensitive information can only be accessed either via the on-site Safe Room or using the online Secure Lab to access the data remotely. The Department for Education, who deposited the data, may be informed about usage.	Department for Education

Data source	Timescale	Method	Geographical coverage	Population age range
Annual Population Survey (APS)	2004 to present.	Annualised version of Quarterly Labour Force Survey which collects information from around 90,000 people every three months about their labour market status, education and qualifications, demographics etc. Around 40% of individuals in each quarter will be included in the APS dataset.	UK	15 years and over.
Annual Survey of Hours and Earnings (ASHE); formerly New Earnings Survey (NES).	ASHE from 2004 onwards, New Earnings Survey (NES) from 1970's to 2003.	Longitudinal/panel/cohort, Postal survey - completed by the individual's employers. Annual waves.	United Kingdom.	One per cent sample of individuals from National Insurance records. Working individuals aged from 16 years residing and working in the UK.
Aspires/Aspires 2	The first part of the Aspires research ran from 2009-2013. Aspires 2 started on 1st February 2014 and will run to 2018.	Aspires: Student surveys (Years 6, 8 & 9), interviews with students and parents (Years 6, 8 & 9). Aspires 2: Student surveys (Years 11 & 13), interviews with students and parents (Years 11 & 13)	England	The first part of Aspires tracked young people from age 10 to 14. Aspires 2 will track the same population from 15 until age 18. Also, parents of young cohort members.
Avon Longitudinal Study of Parents and Children (ALSPAC), Children of the 90s; ALSPAC-G2, Children of the Children of the 90s, COCO90s	ALPAC: 1990 - present. ALSPAC-G2: rolling basis as original youth cohort have children of their own and funded until March 2019	Questionniares, health assessments, DNA & other biological samples. Data collection timepoints vary with age but intervals range from a few months to two years.	Bristol area, UK. Participants were recruited from the former health areas of Southmead DHA, Frenchay DHA, and Bristol & Weston DHA.	Young cohort: fetus to current age (~25-27). Parent/(ALSPAC2 Grandparent) cohort: age at time of pregnancy to current age.
British Cohort Study 1970, BCS70	1970-2005	Longitudinal/panel/cohort. Interviews, surveys, health assessment, clinical records, diaries. Data collection ocurred at intervals of four to ten years.	England, Scotland, Wales	Birth to 42 years and onwards for the young cohort. Information collected about mothers at birth. Survey of teachers and head teachers when child aged 16.

Data source	Timescale	Method	Geographical coverage	Population age range
British Council, Next Generation	First survey: 2011, Latest survey: 2016.	Mixed-methodology approach: desk-based literature review, national survey with a diverse segment of the young population which covers both urban and rural locations, interviews and/or focus group discussions. Data collection interval of five years between first and second wave.	UK, other reports in this series cover Pakistan, Nigeria, Bangladesh, Tanzania, and Ukraine. Research for South Africa, Turkey and Kenya is forthcoming	UK undergraduates aged 19-21
British Household Panel Survey (BHPS), 1991-2009, Understanding Society, United Kingdom Household Longitudinal Study; (UKHLS), 2009-2015	BHPS: 1991-2009￼UKHLS: 2009 to present	Longitudinal/panel/cohort, Face-to-face interview; Telephone interview; Self-completion, paper survey, Computer Assisted Personal Interview (CAPI). Annual waves.	England, Northern Ireland, Scotland,Wales	BHPS: adults aged 16 years and over, household members aged 11-15 from Wave 4;￼UKHLS: adults aged 16 years and over, household members aged 10-15;
British Social Attitudes Survey (BSA)	1983 to present.	Repeated cross-sectional study. One part administered Face-to-face interview; one part administered as Self-completion. The interview is conducted by Computer-Assisted Personal Interviewing (CAPI). Annual waves.	Great Britain: England, Wales and Scotland - excluding the 'crofting counties' north of the Caledonian Canal.￼A similar survey in Northern Ireland is the Northern Ireland Social Attitudes Survey.	Adults (18 and over) living in private households in Great Britain (excluding the 'crofting counties' north of the Caledonian Canal).
Business and Community Interaction survey (HE-BCI)	1999 to present.	The data is collected annually by the Higher Education Statistics Agency (HESA) as part of the Finance Statistics Return (FSR). It comprises a qualitative questionnaire, and quantitative data submission.	UK	Not applicable - data collected at intitutional level, not individual level.
Census: Population census, and Microdata	Aggregated data: Every ten years since 1801, apart from 1941.￼Microdata: datasets are available for 1991, 2001 and 2011. The 1991 and 2001 datasets are referred to as Samples of Anonymised Records (SARS).	Questionnaire; paper survey, online survey.￼The Microdata consists of random sample of 10 per cent of people or households in the census output database for England and Wales. The household and individual files do not overlap and cannot be combined. Data is collected every ten years.	The Census occurs simultaneously in all parts of the UK. In England and Wales, the Office for National Statistics (ONS) is the responsible body. In Scotland, it is the National Records of Scotland (NRS), and in Northern Ireland, the Northern Ireland Statistics and Research Agency (NISRA).	All ages

Data source	Timescale	Method	Geographical coverage	Population age range
Census: Neighbourhood statistics	NeSS includes 2001 and 2011 Census data.	The census data was collected using a paper survey with the option to complete an online survey. Other government data sources, such as ONS and DfE may include administrative data or national surveys. Data is collected at ten year intervals.	NeSS covers England and Wales only. Data for Scotland is available from Scottish Neighbourhood Statistics (SNS). Data for Northern Ireland is available from the Northern Ireland Neighbourhood Information Service (NINIS).	All ages
Census: ONS Longitudinal Study	The sample was initiated at the time of the 1971 Census, and was updated at the 1981,1991 2001 and 2011 Censuses.	Longitudinal, panel, cohort. Self-completion; online option. The sample comprises people born on one of four selected dates of birth and therefore comprises around 1% of the total population. New LS members enter the study through birth and immigration and existing members leave through emigration or death, however their data is retained.Thus, the LS represents a continuous sample of the population of England and Wales, rather than a sample taken at one time point only.	England and Wales. A similar study in Scotland is the Scottish Longitudinal Study (SLS). A longitudinal study in Northern Ireland links Census data with health records: the Northern Ireland Longitudinal Study (NILS)	All ages
Citizenship Education Longitudinal Survey (CELS) ; also incorporates Citizens in Transition: Civic Engagement among Young People, CiT;	2001-2014.	Case study visits to schools: student discussion groups, interviews with teachers and head teachers; regular longitudinal surveys with young people, teachers and school leaders. A cross-sectional survey with young people outside of the cohort. Ongoing literature review. Data collection occurred at intervals of one to three years.	England	Cohort sample: 11 (year 7) - 23 years; Cross-section sample: 12 (year 8) - 25 years; Teachers and head teachers: adults
Effective Pre-School, Primary and Secondary Education project (EPPSE)	1997 - 2013	Face-to-face interview; Postal survey; Educational measurements; Observation; Physical measurements. Data collection occurred at intervals of one to three years.	England A similar study in Northern Ireland is the Effective Pre-school Provision in Northern Ireland (EPPNI)	3 years to 16 years old. Parents and teachers of young cohort.
Employer Perspectives Survey	2007 to present.	Telephone interview. Biannual waves from 2010.	United Kingdom	Not applicable - data collected at workplace level, not individual level.

Data source	Timescale	Method	Geographical coverage	Population age range
ESPON Database	Varies with datasets. For example, 1998 - 2001.	The data included in the ESPON Database mainly comes from European institutions such as EUROSTAT and EEA, and from all ESPON original research projects.	The European Union, Switzerland, Norway, Iceland and Liechtenstein (ESPON space).	Education: birth to adult; employment: 15-65 or older
European Longitudinal Study of Parents and Children (ELSPAC)	Starting with pregnancies/births in 1991-1995 to the present.	Questionniares, health assessments, DNA & other biological samples. Surveys are conducted at: the 18th-20th week of pregnancy, at delivery, in the sixth week, in the sixth and eighteenth month of childrens' age, at three, five, seven, eleven years of age, fifteen years of age at least till the 18th year of age. Participating countries have been encouraged to continue the study until age 21 years.	Participating countries: Great Britain (ALSPAC), Isle of Man, the Czech Republic, Slovakia, Russia, Ukraine, part of Croatia and Estonia, and initially also in Greece and Spain.	Young cohort: fetus to current age (~25-27). Parent/(ALSPAC2 Grandparent) cohort: age at time of pregnancy to current age.
eurostat	Varies for different topics and data sources. For example LFS main indicators are provided for ten years [2006-2015], whereas participation in education is presented for three years [2013-2015].	National administrative data, national survey data including Labour Force Survey (LFS) and European Labour Force Survey (ELFS).	UK, EU countries and also some others.	Education data covers all stages of education from early to higher. Labour statistics include individuals aged 15 years and above.
Families and Children Study (FACS); formerly known as the Survey of Low Income Families (SOLIF)	1999-2009. Low income families from 1999 (wave 1) onwards, high income families from 2001 (wave 3)	Longitudinal/panel/cohort, Face-to-face interview; Self-completion. Annual waves.	England, Scotland,Wales	Parents - defined as anyone over the age of 15 years who had parental custody of either a child aged 16 years or less, or 18 years or less and in full-time education. Children aged 11-15 years.

Data source	Timescale	Method	Geographical coverage	Population age range
Futuretrack Study	2005/2006 cohort followed for six years.	Longitudinal survey which followed a single cohort of 2005/2006 HE applicants. Surveyed four times in six years.	United Kingdom	HE students aged 17 and over.
HESA Student Record	1994/95 - present	Student records from all UK-based, publically funded HEP's. Annual waves.	United Kingdom	Students in higher education, typically aged 16 and over.
HESA Destinations of Leavers from Higher Education (DLHE) survey, initial or early DLHE	Annual waves: data has been collected as DLHE from 2002/03 but discontinuities in data collection means comparisons over time can be limited. For example the survey was redesigned in 2011/12 to collect richer information. Data will continue to be collected until the graduating cohort of 2016/2017 - with the last surveys taking place in April 2017 and January 2018 (with data published in June 2018). HESA is undertaking a transformation programme 'Data Futures' to support a more modern and efficient approach to collect and deliver data, and this involves a new model for the DLHE (both initial and longitudinal surveys, see below). The new survey has been labelled as NewDLHE, and this will cover those graduating from 2017/18 onwards with surveys in December 2018, and then March, June and September 2019. The NewDLHE will cover the same population as DLHE adn will be a census of all graduates, but will be administered centrally by a survey contractor and will explore destinations 15 months after graduating.	Survey. The survey is sent to graduates by their Higher Education Institution. This data is then submitted to HESA. In recent years this has been an online survey supplemented by telephone surveys. Surveys have taken place twice a year to try to account for different leaving dates within the academic year - giving two potential census dates one in April and the and one in January (this was 14 April 2016 and 12 January 2017 for 2015/16 graduating cohort).	Graduates from publically funded HEIs (plus the University of Buckingham) across the whole of the United Kingdom (across all devolved administrations). But includes individuals from all domiciles not just UK domiciled.	Graduates from higher education typically aged 16 or over.

Data source	Timescale	Method	Geographical coverage	Population age range
HESA Destinations of Leavers from Higher Education (DLHE) Longitudinal survey	The first cohort tracked were those graduating in 2002/03, the latest data are available for those who graduated in 2010/11 but data are currently being captured for the 2012/13 graduating cohort.This is likely to be the last cohort tracked with the current approach as HESA is undertaking a transformation programme 'Data Futures' to support a more modern and efficient approach to collect and deliver data. This involves a new model for the DLHE (both initial and longitudinal surveys,). NewDLHE: The new survey has been labelled as NewDLHE, and will cover those graduating from 2017/18 onwards with surveys in December 2018, and then March, June and September 2019. The NewDLHE will cover the same population as early DLHE and will be a census of all graduates, but will be administered centrally by a survey contractor and will explore destinations 15 months after graduating.	Survey. The survey is sent to a sample of respondents to the initial or early DLHE sample. The sample are selected centrally by HESA and allow for some groups to be boosted or over sampled to ensure sufficient numbers for analysis. A second sample can also respond to LDLHE (although only online) and these include all graduates with a current email address or mobile telephone number who responded to the initial DLHE but have not been specifically sampled. The LDLHE survey is administered centrally by a survey contractor (currently IFF Research) but contact details are provided by the relevant HE provider. Graduates are surveyed approximately 3.5 years after leaving their university or college (for the 2010/11 cohort the census date was 24th November 2014). Sampled individuals can respond via an online survey questionnaire or by telephone.	Graduates from publically funded HEIs (plus the University of Buckingham) across the whole of the United Kingdom (across all devolved administrations).	Graduates from higher education typically aged aged 21 or over.
Individualised Learner Record, ILR	2003/04 to present.	2013-onwards: Administrative data submitted annually by organisations providing state funded Further Education. 2003/04-2013: Form returns filled by The Data Service and sent via an electronic system to the Information Authority.	England See also: Northern Ireland - Further Education Regulated Enrolments Scotland - Infact database Wales: Lifelong Learning Wales Record (LLWR)	Those participating in Further Education, typically 16+ years. earners who are 'subcontracted' in from a school or higher education institution (ie learners attending a school or higher education institution who undertake a particular course at an FE institution) are not included in the individual level returns.
Labour Force Survey (LFS) Longitudinal Datasets	Quarterly	Subset of the Quarterly Labour Force Survey which collects information from around 90,000 people every three months about their labour market status, education and qualifications, demographics etc. Respondents in the LFS are interviewed five times at quarterly intervals.	United Kingdom	15 years and over.

Data source	Timescale	Method	Geographical coverage	Population age range
Life Opportunities Survey (LOS)	2009-2014	Longitudinal/panel/cohort. Face-to-face/telephone interview. Annual surveys.	Great Britain	People aged 16 years and over. Parents or guardians were also asked to provide some key data about children, aged 11 to 15. There were three popupation groups: a group with disabilities, a comparison group of with no disability, and a larger group without disabilities, monitored for the onset of impairment over time.
Local authority interactive tool (LAIT)	Data availability varies by data variable. The data release from 28 February 2017 includes data from 2007 to 2016 for some variables. Data is updated on an ongoing basis with updates occurring approximately once a month on average in 2016.	Interactive Excel spreadsheet tool containing linked administrative data which can be used to view numbers and some basic percentages for a selected local authority and to compare it with other local authorities or at a regional level.	England	Birth to 19 years
Longitudinal Education Outcomes (LEO)	Graduate outcomes dataset: tracks those graduating in 2003/04 through to 2012/13 cohorts, and aims to track destinations up to the year 2023. Key stage 5 outcomes dataset: tracks the 2012/13 cohort into the following year (ie destinations in 2013/14).	Linking of administrative data sources to monitor destinations over time. Currently links NPD, HESA Student Record, ILR, HMRC employment data (using the National Insurance and PAYE system), and DWP National Benefits Database data. There are also plans to supplement data for earlier cohorts with data from the Work and Pensions Longitudinal Study (WPLS),and potentially for LEO to be linked to the new DLHE survey in 2020 (see DLHE entry). Data are matched using an algorithm that takes into account: National Insurance number, name (forename and surname), date of birth, postcode and gender. These data are only used for matching. Data for analysis is anonymised. For the graduate outcomes dataset, outcomes are assessed at one, three, five and, in future, ten years after graduation	England	Graduate dataset: Cohorts graduating in years 2003/4 to 2012/13 (although linked data may be available for earlier cohorts 1995 onwards) Key stage 5 dataset: 2012/13 cohort into 2013/14 destinations.

Data source	Timescale	Method	Geographical coverage	Population age range
First Longitudinal Study of Young People in England (LSYPE1); Next Steps.	2004-2010	Longitudinal, panel, cohort. Face-to-face interview; Telephone interview; Self-completion. annual, Waves 1-7	England	13-14 years old (year 9) to 19-20 years old. For Waves 1-4, at least one of the young person's parents/guardians/carers were also interviewed.
Second Longitudinal Study of Young People in England (LSYPE2); Our Future.	April 2013 - September 2013 (Wave 1); intention to continue for 7 annual waves in total.	Longitudinal, panel, cohort. Face-to-face interview; Self-completion. 7 annual waves are planned.	England	Young people aged 13-14 years (Year 9) annually through to the age of 20 (seven waves); and their parents.
Millennium Cohort Study (MCS); Child of the New Century	2000 - 2018.	Longitudinal/panel/cohort . Waves occur every two to three years. Five waves have been conducted to date: MCS1 (age 9 months), MCS2 (age 3 years), MCS3 (age 5 years), MCS4 (age 7 years) and MCS5 (age 11 years). MCS6 (age 14) is currently in progress and MCS7 (age 17) is planned for 2018.	United Kingdom. Teacher interviews in Wave 5 only conducted in England and Wales.	Parents of children from birth to eleven years (in their last year of primary school). Children aged eleven years (wave 5). Some primary school teachers of children in England and Wales.
National Pupil Database (NPD)	2002 to present.	This is an administration dataset with returns made by schools, education providers, Local Authorities, and qualification awarding bodies. It includes all pupils in the English state maintained sector. This is a live database. Data collection is ongoing with data added regularly. Different sources feed into NPD in different months.	England. See also: Northern Ireland - Schools Census Scotland - Pupil Census Wales - Schools Census	Young people in education from 0 to 19 and over.

Data source	Timescale	Method	Geographical coverage	Population age range
OECD	Data presented on the website is for the current and previous year, e.g. 2016 and 2015. Some historical data is available for some variables, e.g. new entrants to education from 1997 onwards.	National administrative data, national survey data including Labour Force Survey (LFS) and European Labour Force Survey (ELFS).	International	Population in education or employment.
Opinions and Lifestyle Survey (OPN), formerly Office for National Statistics (ONS) Opinions Survey or ONS Omnibus Survey. It became the Opinions and Lifestyle Survey in April 2012 following the merger of some elements from the General Lifestyle Survey (GLF), which closed in January 2012.	October 1990 to present.	Face-to-face interview. Collection varies between monthly and eight cycles per year. Since April 2014 survey data is not collected in March, June, September or December.	Great Britain, excluding the Isles of Scilly and the Scottish Highlands and Islands	Adults, aged 16 or over, living in private households in Great Britain.
School Performance Tables	1991/92 to present. The tables are fed by live data with ongoing updates. Earlier years are available as data downloads.	This draws in administration information from the schools census (on number and profile of pupils), local authorities (on attainment, assessment and progress at KS2), and qualification awarding bodies (attainment, assessment at KS4 and KS5). Destinations data draws from the linked NPD data.	England See also: Northern Ireland - Annual Examination Results (AER), not available at individual school level. Scotland - Curriculum for Excellence Levels Wales - My Local School contains performance data using the National School Categorisation System.	Not applicable - data collected at school/college level including primary, secondary and 16 to 18 schools and colleges.
Universities and College Admissions Service (UCAS)	1992 (UCAS formed) to present	Application data collection, annual.	United Kingdom	Applicants to HE typically aged 17 years and over.

Data source	Timescale	Method	Geographical coverage	Population age range
UCAS Strobe	1992 (UCAS formed) to present.	Administrative data from the UCAS application process. Data collected annually.	UK	Applicants to higher education typically aged 17 or over.
Young People's Social Attitudes (YPSA)	One-off surveys in 2003, 1998 and 1994.	Cross-sectional (one-time) study. Face-to-face interview. Data collected at intervals of four to five years.	Great Britain: England, Wales and Scotland - excluding the 'crofting counties' north of the Caledonian Canal. A similar survey in Northern Ireland is the Young Persons' Behaviour and Attitudes Survey(YPBAS)	Young people aged 12-19 living in the households of BSA respondents
Youth Cohort Study (YCS)	1985 to 2010.	Longitudinal/panel/cohort. Online, telephone and face-to-face interviewing,. Email survey; Web-based online survey. New cohorts start at 1 to 3 year intervals and are then followed through three annual waves. The most recent cohort is Cohort thirteen and its fieldwork covers 2007-10.	Cohorts One to Twelve cover England and Wales but from Cohort Thirteen (those eligible to leave in 2005-6), data covers England only. A similar study in Scotland is the Scottish Young People's Survey (SYPS).	Young people who reached minimum school leaving age in a given school year (aged around 16) who are followed for four years.For example, those eligible to leave school in 2005-06 were first interviewed in 2007, and then in 2008,2009,2010.

Data source	Relevance	Limitations of data source
Annual Population Survey (APS)	Relevant topics include: economic activity status ie employee, full-time student etc; characteristics of employment eg sector, occupation, ft/pt etc.; qualifications; education or training; apprenticeship; demographic characteristics eg gender, age, region etc. The longitudinal element allows an examination of the sizes and characteristics of the flows of young people from education and employment/unemployment/inactivity between the different quarters	The data are not longitudinal, only cross-sectional, and there very few and limited retrospective questions ie circumstances one year ago, therefore the potential to investigate the education to work transition is very limited. However the large sample size means that the labour market circumstance of individuals by single year of age (eg 16 year olds, 17 year olds etc.) can be investigated.
Annual Survey of Hours and Earnings (ASHE); formerly New Earnings Survey (NES).	Relevant information for the employee include: NI number, employment dates, contract type (permanent), apprentice, pay, bonuses/incentives, overtime, pensions, pension contributions, salary sacrifice, benefits in kind, paid leave entitlements, pay agreements, postcode (employee,employer), job title and description, occupation and industrial classifications. There are also variables for age, gender and full/part-time status. As data is collected by the employer, there are variables relating to the employer such as size and legal status (e.g. public company).	This dataset provides very detailed information about the pay and benefits of a cohort of adults in employment. It does not include individuals who are self-employed, or not paid through PAYE. There is very limited information about the employee's role, and no information about qualifications or education. Within the dataset, it is possible to track an individual across job roles but not necessarily the initial transition from education to work (unless an apprentice). NI number is collected which means that individual data could be linked to some other national datasets.
Aspires/Aspires 2	Relevant topics include: young person demographic information, attitudes, influence of family and school, school attended, education choices and attainment, careers aspirations and careers guidance, social identities	The data present a rich source of information regarding career aspirations of young people in England, tracking them from ages 10 to 18. However, it's geographical scope is limited to England and does not explore experiences of young people throughout the UK. It does not explore outcomes for young people after compulsory education until 18, although if the necessary identifiers are present it may be possible to link to other datasets such as HESA, ILR and track these individuals in future for their longer term employment/learning outcomes. While there is discussion of careers guidance, the focus of this study is on science and science careers so the data may be of weaker relevance to other areas of learning or employment, although it it likely that the research would include some more general questions regarding aspirations and careers.
Avon Longitudinal Study of Parents and Children (ALSPAC), Children of the 90s; ALSPAC-G2, Children of the Children of the 90s, COCO90s	Relevant topics include: information about the child's family and household, some education outcomes, some literacy and ongoing mathematics tests, child's attitude and behaviour in school, setting and streaming, school environment and teacher, very rich information on child's health and development including DNA samples.	The data focuses on health and school-based educational outcomes but could potentially be linked to other datasets. The study covers a samll geographical area as the sample is drawn from the Bristol area of England. The dataset has very rich information around the health, family environment and behaviour of the young person. It explores the child's behaviour in school, setting and streaming, and mathematical skills, as well as incorporating some data on educational outcomes from NPD data. However, information about careers interventions and guidance, or aspirations/expectations for the future were not collected. Data collection is still ongoing with the original youth cohorts as adults and with their children which could allow for observing impacts across generations with a very detailed picture of the parents. However, this data only covers a local population specific to the area around one city in England although comparisons have been undertaken to compare how representative the Avon sample is of the overall UK population. The Principle Investigators are exploring the potential to link the data to information about benefits, earnings and employment data.
British Cohort Study 1970, BCS70	Topics include: health (mother and child), education including literacy and numeracy, social and economic circumstances. The wave at age 16 includes a survey with questions about attitudes to work, expectations and career preferences; the teacher's survey includes O-level/CSE predictions. The wave at age 21 incuded questions on: employment and education histories since age 16, qualifications, training,unemployment, reading and writing behaviour, literacy and numeracy self-appraisal and assessments, income, attitudes to employment, education, literacy and numeracy, and self efficacy.	The dataset provides a detailed picture of health, family circumstances and education development for the young cohort. Attitudes to work and expectations are explored at age 16 and later data may be linked to DWP/HMRC data where permission is given. As the cohort is born in 1970, this data may be less relevant as education and qualifications have changed significantly during this time, e.g. O-levels/CSEs dichotomy, and the types of employment have also changed with technological development and increased globalisation. However, it may be useful to compare young people's attitudes/expectations around work, and to explore how/whether long term employment outcomes relate to attitudes at age 16.

Data source	Relevance	Limitations of data source
British Council, Next Generation	The survey and focus group include some questions around experiencs of working or volunteering in another country, some attitudes to working/volunteering/studying in another country, and some questions on where they have heard about opportunities to work/volunteer/study abroad. Respondent gender and ethnicity is recorded. The first survey undertaken in a country records a baseline for benchmarking, and later reports are planned to explore selected topics and themes in more detail.	The age range of respondents in this survey is quite narrow (19-21) and it's a specific segment of this group, ie those participating in undergraduate degrees. There are some limited questions on attitudes to work and volunteering in an international context but mostly the research focuses on engaging with international opportunities and global issues. There is limited demographic information recorded, primarily: degree subject, higher education institution, gender, age and ethnicity.
British Household Panel Survey (BHPS), 1991-2009, Understanding Society, United Kingdom Household Longitudinal Study; (UKHLS), 2009-2015	BHPS: Relevant topics include: neighbourhood; individual demographics; residential mobility; health and caring; current employment and earnings; employment changes over the past year; lifetime childbirth, marital and relationship history; employment status history; values and opinions; household finances and organisation. UKHLS: Relevant topics include: demographics, family background, ethnicity and language use; migration, health, disability and caring; current employment and earnings; employment status ; parenting and childcare arrangements; family networks; benefit payments; household finances; environmental behaviours; consents to administrative data linkage. Young cohort topics include: computer and technology use, family support, sibling relationships, feelings about areas of life, the Strengths and Difficulties Questionnaire (SDQ), and aspirations. Individuals are asked about their educational aspirations, their first job after completing education, and their plans for the future. Young adults survey includes question about which careers advice they have received if any.	This dataset covers individuals across the UK and collects very detailed information about employment, levels of education, income, aspirations, demographics, family and household, health and other activities. It can also be linked to DfE and NHS data. There is a question for young people asking if they have received any advice about careers but it does not ask about other careers interventions, e.g. work experience.
British Social Attitudes Survey (BSA)	The BSA includes a set of core questions which occur in most of the survey waves, and then the remaining questions and topics change across waves of the survey. Sometimes a topic or series of questions is repeated across several waves to observe trends and some questions or topics are one-off instances. Relevant topics from the core questions include demographic information including: gender, age, health & disability, children, marital status, ethnicity, national identity, employment status, income & benefits, last employment, self-employed/employee, how many employees, employer sector, full-time/part-time, hours worked, age left full-time education, qualifcations gained, household members, owning or renting. Core derived variables include standard occupational classification, and socio-economic indicators such as NS-SEC, Registrar General's Social Class, and Socio-Economic Group. Other core questions mostly pertain to political attitudes. Among the changing questions, there may be some that are relevant to CEC's interest areas and organisations can fund a question or questions to be added to waves of the survey by contacting NatCen. For example, the 2015 BSA included the 'work orientations' module from the International Social Survey Programme (ISSP) which includes questions on attitudes to work, worklife balance and experiences of work.	This questionnaire provides a snapshot sample of the general population rather than tracking individuals over time. While there are some core questions asked across most waves of the survey providing personal characteristics information such as employment status, income and last qualification, the topics vary widely and focus on attitudes to social or political issues. Also, this survey covers most of Great Britain but excludes Northern Ireland and parts of Scotland.
Business and Community Interaction survey (HE-BCI)	HE-BCI investigates the interactions of Higher Education Providers with businesses and their local communities. This mostly focuses on researrch undertaken with public funding (not academic), services provided to business and the local community, participation in regeneration programmes, Intellectual Property (IP) activities, and public events for the community. Survey topics relevant to young people's careers and transitions included: one question about HEPs' involvement with student and graduate enterprises and start-up, a question exploring services such as 'incubator' faciliaties for student or graduate start-ups.	HE-BCI records information at an insitutional level and so cannot be used to monitor transitions, attainment or attitudes information for individuals. There is very limited information relevant to careers guidance or support provided for students and young people, there are two questions which include information about whether the institution is involved with student and graduate enterprises/start-up, and whether the insitution provides 'incubator' facilities for student/graduate start-ups.
Census: Population census, and Microdata	Relevant topics include: gender, ethnicity, economic activity/employment status, highest level of qualification, occupation. The census is the most complete source of information about the population. As the same questions are asked of all households at the same point in time, it is partiuclarly useful for making comparisons across different parts of the country.	Aggregate data: Data is presented at an aggregate level so it is not possible to track individuals through the dataset. The data does not capture information about careers advice or interventions, and it does not collect attitudinal information such as future aspirations. As the census is conducted at ten year intervals, this means that it is a relatively course longitudinal measure. Microdata: Topics covered are the same as with the aggregated data. Individuals and households can be identified but these datasets are one-off 'snapshots' at a point in time. In order to follow individuals over time across censuses, researchers would need to access the Census Longitudinal Study dataset.

Data source	Relevance	Limitations of data source
Census: Neighbourhood statistics	The NeSS data includes Census topics such as gender, ethnicity, economic activity/employment status, highest level of qualification, and occupation. In addition to this, NeSS incorporates modules from national-level sources such as educational attainment from DfE and information regarding health from the NHS Information Centre for Health and Social Care.	Data is presented at an aggregate level so it is not possible to track individuals through the dataset. However, it does provide a range of area category variables such as educational authority and down to areas with 250 people which allows for quite detailed pictures of areas of England or Wales. The data does not capture information about careers advice or interventions, and it does not collect attitudinal information such as future aspirations. As the census is conducted at ten year intervals, this means that it is a relatively course longitudinal measure.
Census: ONS Longitudinal Study	The LS data comprises completed census returns from 1971 onwards, linked with life event data from various sources. Data is available for both LS members and the people they live with, and includes: age, sex, marital status; family, household or communal establishment type; housing, including tenure, rooms and amenities; country of birth and (in 1971) parent's country of birth; ethnicity (1991 and 2001); educational qualifications; economic activity; occupation and social class.	The Census longitudinal data gives researchers the opportunity to track basic employment and qualification details for individuals over a long timescale. However, as there is a ten year interval between data collection waves, this is a very coarse measure and may not be indicative of the general pattern of employment over the previous ten years. There is no information collected about experiences of careers advice and guidance, experiences of careers interventions and activities, and aspirations or attitudes to future work or study.
Citizenship Education Longitudinal Survey (CELS) ; also incorporates Citizens in Transition: Civic Engagement among Young People, CiT;	This study has limited relevance but provides an opportunity to explore some basic education/employment outcomes with regards to interests and activities, such as volunteering and activism as well as young people's perceptions as to whether volunteering will help make them more employable. The main focus of this study explores experiences and views about citizenship in classrooms, schools, and the wider community. The research collects some information on parent work and education, school factors, and quite detailed information on interests and activities, especially volunteering or political activism. The longitudinal survey element captures basic information on transitions from compulsory education to employment or education outcomes as a young adult; however, it does not capture information about grades/performance. The cross-sectional surveys ask some questions about qualifications obtained and current economic activities. Outcomes measured for students include: political literacy, political efficacy, voluntary action, social norms and values, identification with institutions, identification with communities, trust in other people and in government, and a sense of rights and obligations. Data is collected for England only.	The main focus of this research is on the impact of citizenship education on the attitudes and behaviours of a cohort of young people in England. It also captures some basic information on family, school, qualifications obtained, and educational/employment outcomes up to age 23 years.
Effective Pre-School, Primary and Secondary Education project (EPPSE)	Relevant topics include: child development in terms of cognitive, educational and social development; family characteristics; pre-school characteristics; primary school characteristics; secondary school characteristics. Six months after completing compulsory education at then 16 years, young participants were sent a 'Life After Year 11' questionnaire which includes questions about: their post-16 destinations, sources of advice and guidance for their post-16 choices, whether they felt they would be likely to go to university and why, and careers aspirations.	The dataset provides information regarding educational outcomes and influences on young people from pre-school until six months after year 11 but it does not capture longer term outcomes. However, as it already links to the NPD, there may be a way to link this data with other datasets such as HESA data. Also, the data explores sources of informaton/guidance for post-16 destination choices but does not explore careers interventions such as work experience.
Employer Perspectives Survey	Relevant topics include: Recruitment processes/service/channels, including young person/leaver recruitment; importance of factors looked for in candidates e.g experience/qualifications; school leaver skills lacking/preparedness; work placement scheme details, post-placement prospects; reasons for offering placements/internships; reasons for not/difficulties when/ how to encourage engaging with educational institutions; problems/ help with training related issues; use/awareness of initiatives; how training through educational institutions is funded; benefits of employer training; apprenticeship provision details/types/knowledge/ plans; reasons for not providing apprenticeships; national occupational standards.	EPS includes information regarding recruitment practices, attitudes, and prospects offered by employers, which gives some insight into this transition stage for young people. However it does not include demographic information of current and past work placement/interns, graduates and apprentices so it is not possible to build a picture of how people are progressing through their career. The data for the 2016 Employer Perspectives Survey collected in November 2016 has not yet been released. UKCES ceased operation in March 2017 so on-going delivery of this research and its outputs will be transferred to another government body.

Data source	Relevance	Limitations of data source
ESPON Database	The theme 'Social Affairs' includes the sub-themes 'Education' and 'Labour Market'. The sub-theme 'Education' includes indicators: early school leavers, enrolment in stages of education, e.g. tertiary. The sub-theme 'Labour Market' includes indicators: labour force participation variables by age and gender, long term unemployment, and unemployment rate. The theme 'Economy, finance and trade' includes the sub-theme 'Employment' which includes indicators such as: economically active population, employment by age group and gender, employment by economic activity, employment rate, overall employment rate, persons in employment/unemployment, typology of gender differences on the labour market.	Data deposited from ESPON research projects tends to focus on specific issues for specific time periods so may be more suited to queries regarding specific areas of interest. However, the ESPON Database includes more general, long-running data from sources such as Eurostat. Data is is presented at a country or NUTS regional level so it is not possible to follow individuals through transitions or use personal identifiers to link to other data sources. The database did not appear to include variables recording transitions from education to work.
European Longitudinal Study of Parents and Children (ELSPAC)	Relevant topics may include: information about the child's family and household, some education outcomes, some literacy and ongoing mathematics tests, child's attitude and behaviour in school, setting and streaming, school environment and teacher, very rich information on child's health and development including DNA samples.	The research is carried out and held separately by each participating country. It appears that the countries use a shared set of questionnaires but it is possible these may have been adapted to national circumstances. Assuming the same questionniares are used then the main focus of the studies are on health and education outcomes until age 18 although countries have been encouraged to continue the research until age 21 or further. If so, then it is possible that Indivdiual country level data may capture some information about further/higher education and employment otucomes. The questionnaires do not collect information about careers interventions or guidance, or about future aspirations or expectations.
eurostat	The theme 'Population and Social Conditions' includes relevant topic areas: Education and Training, Labour Market, and Youth. The topic 'Labour market' includes data from the Labour Force Survey such as: employment & activity, self-employment, part-time/full-time employment, temporary employment, second jobs, working time, quality of employment, employment growth, duration of working life, average exit age from labour force, unemployment, long-term unemployment, supplementary indicators, jobless households, regional data, and labour market transitions shows individuals entering employment from unemployment. There is also information on earnings. In 2009, the LFS included an ad hoc module on young people in the labour market including: average when age leaving formal education, educational attainment for those who left, and employment rate by number of years after leaving education. In 2000, a module was included on transition from school to working life which included: transition from education to work by months after leaving education, occupational status, proportion employed in service sector, labour outcomes by educational attainment and social origins, parent educational attainment level, and job mismatch. The topic 'Education & training' includes: participation in lower secondary, upper secondary, post secondary, vocational upper secondary and post-secondary, tertiary education, ISCED11 education levels, and new entrants. Information on transitions from education to work includes: NEET, early leavers from education, labour status if young people by years since completion of highest level of education, and entry of young people into the labour market. There is also information about education outcomes. The topic 'Youth' includes: Youth Education & training which includes educational attainment, early leavers from education & training, tertiary education attainment level, and upper secondary attainment level. The topic 'Youth employment' includes: employment, self-employment, part-time employment, unemployment, and NEET. The topic 'Youth volunteering' includes participation in voluntary activities.	This data source can be used to identify trends at country level for the UK and a group of other mostly european countries; Microdata containing individual level data is also available and so tracking individuals over time through transitions may be possible if indivdiuals are given unique personal identifiers. Eurostat holds some quite detailed information on employment and education including NEET, labour status since completion of highest level of education for young people and ISCED11 education levels. Ad hoc modules in 2000 and 2009 focused on young people and their transitions. In 2009, a module focused on young people in the labour market including average age, educational attainment, and employment rate by number of years after leaving education. A 2000 module focusing on transitions from school to working life may be a bit old but data such as entering work by months after completing education could be compared with more recent data, e.g. the 2009 module, to identify any differences which could suggest possible trends.
Families and Children Study (FACS); formerly known as the Survey of Low Income Families (SOLIF)	Relevant topics from the parents' survey include: household characteristics, children's schooling, behaviour and childcare provision, use of local services, education and training, employment, family income, in-work support through the use of Working Families' Tax Credits (and its replacement tax credit system), receipt of benefits, child maintenance, money management and savings, housing, material deprivation, and social capital. Relevant topics from the children's survey include: leisure time activities, computer access, social participation, sport and organised activities, use of local amenities and attitudes to neighbourhood, self-esteem, health and happiness, attitudes to school and schoolwork, relationship with parents, and income	The dataset provides detailed information on family circumstances and activies for a ten year period, especially for low income families, and with regards to benefits and income received. The focus of the study is on parents but outcomes for children are examined with young people aged 11-15 responding to surveys exploring their attitudes and activities. The studies do not specifically map the transition from compulsory schooling to work/further education but this information may be captured in some cases. Permissions have been sought to link the dataset to DWP and NPD data. Children are asked what they would like to do when compulsory schooling ends at age 16 but they are not asked about careers advice or interventions, although parents are asked about their aspirations for their child/children.

Data source	Relevance	Limitations of data source
Futuretrack Study	Relevant topics organised by wave: Wave one (time of UCAS application): Demographic information; secondary education; attitudes towards, influence and reasons for choosing/withdrawing/changing their HE/course/HEP; family interaction with HE; attitudes towards student funding/debt Available info regarding the following: career implications of post-16 exam subject choices; relationship between course and employment options; classroom career and life planning; HE alternatives; careers guidance; career opportunities from employers; careers fairs. Wave two (Students in first year of HE or withdrawn/deferred): demographic information; attitudes towards current HE course and HEP, including career prospects; paid work related to career plans; employment status at HEP; attitudes towards withdrawal from HE; reason/attitudes/basis regarding paid/unpaid work during HE; funding interventions; experience of careers and employment services organised by HEP, family and employer; family SEC. Wave three: (Conducted with 3rd year final year undergraduates as they approached the end of their courses in summer 2009): demographic information; individual educational information; experience/basis/attitudes regarding course work placement and work experience relevant to course; employment experience; percieved skill competency; expected degree class; employment market knowledge/expectation/attitudes; careers service information; employer interaction; post HE plans; seeking employment, such as applications, industry interests and offers Wave four: individual educational information; reflective attitudes towards HE; main activity history/employment status, paid or unpaid; employer details; in-job training; required qualifications for job; reasons for multiple occupations; current job type, satisfaction, security, earnings, progression; perceived skills; attitudes towards course preparation for employment.	The data tracks a single cohort over 6 years through their HE and any employment or further study/training destinations. The data follows one cohort and data collection was conducted between 2005 - 2013 so the information may be slightly out-of-date and limited when it comes to tracking the impact of recent interventions/policy. For example, tuition fees increased to £9,000 from 2010.
HESA Student Record	Relevant topics include:individual educational information; learner employment status; learner delivery work placement; apprenticeship engagement, including engagement outcome and withdrawal reasons; qualifications awarded; parental education; socio-economic classification; care leaver / looked after status;	The dataset provides some information regarding engaging with employers during higher education such as employment status and apprenticeship programmes. However, the dataset does not record employment outcomes after higher education courses have been completed. Student record data can be linked to DLHE data to track destinations six months after completion of study. The data does not collect information regarding student experiences, attitudes or plans regarding both HE and employment.
HESA Destinations of Leavers from Higher Education (DLHE) survey, initial or early DLHE	Relevant topics include: activity and most important activity on the census date; attitudes towards how HE course prepared leaver for work, business or further study; reasons for taking up employment. What leaver's are doing six months after study. Employment details: On which basis they previously worked for employer, such as placement/internship; how they found employment; current employment basis; employer's name, location, size and nature of business. Job details: title, duties, pay, payment period; total annual earnings which include other roles; qualification required for job; perceived most important element of qualification to employer. For those progressing to further study, details of HE study: course, institution, qualification sought, study mode, how they are funding further study. For those moving to teaching: whether employed as teacher or seeking a teacher post, type of school, and phase.	The DLHE is largely a census of all final year students studying in the academic year from 1 August to 31 July, but currently excludes those studying at FE level (in HE) as these are covered by the ILR, and those studying mainly overseas and incoming exchange students. As noted, discontinuities in data collection over time mean that comparisons over time may be restricted. Also response rates are much lower for international students and so these data tend not to be reported. This dataset provides rich information on destinations six months after completion (or otherwise leaving) their studies, as well as their attitudes towards how their studies prepared them for this situation. However, it does not include longer-term career goals or attitudes towards their current employment/employer. Furthermore, the data collected regarding how prepared they feel for employment only refers to the course, not the careers service provided by the relevant institution. The data concerning a leaver's previous relationship with an employer and the basis on which they are now employed does not currently include apprenticeship pathways as an individual category (although 'Placement' and 'Internship' categories are available). New DLHE: The DLHE and longitudinal DLHE (see separate note) will be replaced for the graduating cohort 2017/18 onwards with a new survey, known as NewDLHE. It is hoped that data will be available from Jan 2020 (for the full year of surveys plus linked data from the student record). This survey is currently in development but will remain a census survey (although administered centrally by a survey contractor) and cover the same population. Graduates will be surveyed by telephone or online 15 months after completing their studies. There will be 4 surveys per year: Dec, March, June and Sept (the latter is likely to be the largest in number). HE providers will be able to add questions to the surveys of their graduates (from an optional bank or bespoke questions) but some charges may apply; and the survey design will allow for additional surveys of sub-populations of interest (at additional cost). NewDLHE also aims to provide institutions with near real-time access to the data (via interactive dash boards). The survey content will be very similar but will capture career history and further study history, so will capture first destinations as well as current activity on the census date. NewDLHE will also capture graduate voice and include questions on meaningfulness/importance of the activity to the graduate, skills utilisation, and progress towards future goals. It will also gather insights into non-traditional career paths such as setting up a business.

Data source	Relevance	Limitations of data source
HESA Destinations of Leavers from Higher Education (DLHE) Longitudinal survey	Various relevant topics are covered. Employment status; and employment details including place, job title, basis of employment, size of organisation, salary, hours worked per week, when first obtained the job, job combination/reasons for undertaking two roles, whether employer provided training. Employment history since graduating; and total unemployment period. Attitudes towards employment: motivations for taking the job, how job was found, career satisfaction, and whether using skills gained from HE. Additonal study: when first started the course, name and type of HE provider, qualification aim and subject, nature of study, funding source and motivations for undertaking further study. Any additional qualifications gained since graduating. Attitudes towards HE course in terms of preparation for workplace: progress towards career aspirations; whether the course represented value for money; and whether HE experience improved skills (in innovation, problem solving, communication, making work decisions, working with others, taking initiative and responsibility, using ICT, and working with numbers). Attitudes towards the importance of the following variables when seeking employment: subject studied, type of qualification, grade of qualification, evidence skills and competancies, work experience/placement whilst on course, qualification gained after HE, work experience from previous employment.	The LDLHE is a sample survey drawn from respondents to the initial DLHE survey and achieves a response rate of between 40 and 50 per cent. It excludes those not covered by the initial DLHE (ie those studying at FE level and those studying mainly overseas and incoming exchange students) and those who opt-out after the initial DLHE. It only covers every other graduating cohort. This dataset provides rich information on destinations 3.5 years after completion (or otherwise leaving) their studies and so allows time for individuals to settle into their careers. It also captures perceptions of the impact of HE study on work outcomes, and skills used in the workplace. The LDLHE and initial DLHE (see separate note) will be replaced for the graduating cohort 2017/18 onwards with one new survey, known as NewDLHE. It is hoped that data from the new survey will be available from Jan 2020. NewDLHE: This survey is currently in development but will be a census survey (although administered centrally by a survey contractor) and cover the same population as initial DLHE. Graduates will be surveyed by telephone or online 15 months after completing their studies. There will be 4 surveys per year: Dec, March, June and Sept (the latter is likely to be the largest in number). HE providers will be able to add questions to the surveys of their graduates (from an optional bank or bespoke questions) but some charges may apply; and the survey design will allow for additional surveys of sub-populations of interest (at additional cost). NewDLHE also aims to provide institutions with near real-time access to the data (via interactive dash boards). The survey content will be very similar but will capture career history and further study history, so will capture first destinations as well as current activity on the census date. NewDLHE will also capture graduate voice and include questions on meaningfulness/importance of the activity to the graduate, skills utilisation, and progress towards future goals. It will also gather insights into non-traditional career paths such as setting up a business.
Individualised Learner Record, ILR	Relevant information Includes: information about learner characteristics, e.g. sex, age , ethnicity, health, identifiers, e.g. NI number, Unique learner number, funding/financial support, employment, previous qualifications, GCSE maths and English grades, postcode, employment, work experience (presumably placement rather than option?), postcode (current & prior to enrolment), UCAS personal identifer, learner employment (before & during FE), programme of learning, apprenticeship details, grades/qualifcations achieved, employment outcomes (where this is a requirement for the course/funding), and destinations, e.g. employment, volunteering.	This dataset holds detailed information regarding learners in further education including previous qualifications or employment and learning outcomes and destinations, e.g. employment. However, it does not track outcomes over a longer timescale. There is potential to link the information to other data sources through unique identifiers including NI number. Also, information is only available since 2003/4 compared with 1994/5 for higher education data.
Labour Force Survey (LFS) Longitudinal Datasets	Relevant topics include: economic activity status ie employee, full-time student etc; characteristics of employment eg sector, occupation, ft/pt etc.; qualifications; education or training; apprenticeship; demographic characteristics eg gender, age, region etc. The longitudinal element allows an examination of the sizes and characteristics of the flows of young people from education and employment/unemployment/inactivity between the different quarters	The data source contains only factual information about status at time of the interview, it does not cover any historical information about pre-16 status or activities, and does not include any attitude questions. Also sample size of the population of interest, ie those in transition between education and work, may be fairly small and thus only suitable for national analysis with limited scope for cross-tabulations of variables.

Data source	Relevance	Limitations of data source
Life Opportunities Survey (LOS)	The LOS questionnaire has two parts: The Integrated Household Survey (IHS) core module collected basic demographic and other information for all members of the household. The LOS module covered a range of topics and was administered to each adult in the household. Demographic information includes: sex, age, ethnicity, nationality, country of birth, national identity, sexual identity, religion, health & disability, government training scheme including one providing work experience, employment, sector, public/private sector, job title, job role, self employment, managerial role, full-time/part-time, looking for work, education, highest qualifications, current course, age finished full-time education, disability characteristics. Relevant topics in LOS module include: barriers to participating in work/ employment, barriers to accessing learning, work hours, limitations to working, work adjustments, support to help work/finding work, e.g. job coach, work negative impact on health/disability, barriers to participating in voluntary/charity work, physical access issues in workplace/educational institution, use of benefits and pension service, discrimination in the workplace, income.	As respondents for the main survey are aged 16 and above it may be possible for some respondents to explore post-sixteen choices and also transitions after age eighteen for some respondents. It will also be possible to compare outcomes for those with a disability and those with no disability. The survey also includes questions on participation in Government Schemes involving work experience, and whether individuals felt they would have benefitted from a work coach so it is possible to explore to a limited extent if there is any difference in participation or perceived need for careers guidance/interventions. Children aged 11-15 years may go on to enter the main survey but the information collected in the questions about young people in the household were very limited and focus on health and disability.
Local authority interactive tool (LAIT)	LAIT was created to benchmark/compare a local authority with its region or its statistically similar neighbours. Relevant topic areas include: GCSE attainment, A-level attainment, attainment by age 19, proportions of pupils going into further/higher education, 5 NEET and skills gap in the current workforce. It also contains contextual information regarding children's services. Some relevant areas include: Children with SEN, free school meals, school exclusion, youth offending, and dedicated schoosl grant.	LAIT provides information at a local authority level so it is not possible to track individuals over time. It provides information on attainment at GCSE, A-level, proportion of NEET, and qualification level attained at age 19. It also provides some` interesting contextual information such as proportions of free school meals, and local skills gaps. This dataset is not suitable for tracking individuals through transitions but does allow for some comparison of educational attainment and basic level outcomes, e.g. NEET, education/training, at age 19. It also includes groupings for 'statistically similar' local authorities for comparisons. LAIT could be a useful tool for comparing outcomes between local authorities to identify areas where careers interventions may be particularly beneficial or for identifying comparator groups for testing careers interventions.
Longitudinal Education Outcomes (LEO)	Relevant areas available in LEO datasets include: personal characteristics (age, gender, ethnic group, household income, eligibiilty for free school meals); education (schools/colleges/universities attended, courses taken and qualifications achieved); any careers advice and training offered by the government; employment and income; and benefits claimed. More specifically, relevant graduate dataset variables include: ethnicity, sex, age, region of domicile, prior attainment (tariff points), full-time/part-time study, subject of HE degree study, Higher Education Institution (HEI), economic activity (sustained employment, further study or no sustained destinatinon), and earnings (calculated as annualised earnings)for those in sustained employment. Key stage 5 dataset variables include: institution type (e.g. state), gender, free meal status in year 11, special educational needs, learners with learning difficulties/disabilities.	The graduate outcomes experiment dataset provides the opportunity for long-term tracking of outcomes in terms of employment status and income, and currently tracks ten graduate cohorts assessing outcomes at regular intervals over a ten year period. However it does not provide detail on employment characteristics beyond income (eg type of employment, working pattern and work location); it does not capture information as to why individuals are experiencing these outcomes, e.g. career break for children, and perceived value of their education/qualifications to their achieved outcomes; and doesn't capture perceptions of success or satisfaction. It does not cover graduate cohorts in Scotland, Wales or Northern Ireland. Other more specific limitations are that: employment data for individuals earning below tax thresholds or are self employed are not currently captured which has a large impact on arts graduates, and arts focused institutions, and graduates working abroad are not covered; benefits data currently excludes Universal Credit information. Also outcomes are measured during financial years whereas learning activities are measured using academic years and this can be confusing. The key stage 5 cohort currently only examines destinations at one year after key stage 5 but potentially these data could be linked to track a key stage 5 cohort into further destinations. Again, this data does not cover cohorts in Scotland, Wales and Northern Ireland. It is not clear whether these datasets capture information on careers guidance and interventions. They may capture inputs through the benefits system but may not capture inputs during compulsory or further and higher education.

Data source	Relevance	Limitations of data source
First Longitudinal Study of Young People in England (LSYPE1); Next Steps.	Relevant topics in Waves 1-4 include: the young person's family background, parental socio-economic status, personal characteristics, attitudes, experiences and behaviours, attainment in education, parental employment, income and family environment as well as local deprivation, the school(s) the young person attends/has attended, and the young person's post-16 plans. Relevant topics in Waves 5-7 include: demographics, activity history and current activity, jobs and training, qualifications being studied, higher education, attitudes to work and debt, childcare and caring responsibilities, young people Not in Education Employment or Training (NEET), Apprenticeships, information, advice and guidance, risk behaviours, relationships and sexuality, and own children.	While the dataset is very rich in terms of information regarding educational characteristics, personal characteristics, employment and careers activities. It is limited to England and does not include the other UK nations. It tracks young people closely for the majority of the time period of interest to CEC, from 13 to 21 years old. It does not include longer term careers or educational outcomes. However, if permission can be obtained to link this with DWP and BIS data in future then this dataset could be revisited to look at longer term outcomes.
Second Longitudinal Study of Young People in England (LSYPE2); Our Future.	Relevant topics include: young person's family background; parental socio-economic status; personal characteristics; attitudes, experiences and behaviours; parental employment; income and family environment as well as local deprivation; the school(s) the young person attends/has attended;the young person's future plans.	While the dataset is very rich in terms of information regarding educational characteristics, personal characteristics, employment and careers activities. It is limited to England and does not include the other UK nations. It tracks young people closely from 13 to 21 years old, which is the majority of the time period relevant to CEC's activities (12-18 years) . It will not cover longer term careers or educational outcomes. However, for LYSPE1 permission can be obtained to link this with DWP and BIS data in future and if this approach is replicated with the LYSPE2 data then it may be possibile in the future that this dataset could be revisited to look at longer term outcomes .
Millennium Cohort Study (MCS); Child of the New Century	Relevant topics include: information about the child's family and household, education, childcare activities, income, housing and local area, cognitive tests with the child , child's attitude and behaviour in school, setting and streaming	This dataset is part of an ongoing study tracking a cohort from birth to age seventeen. So far, only data for ages up to 11 years is available but data for age 14 is currently being collected and age 17 data will be collected in 2018. Where individuals have given consent, this dataset could be linked to NPD data for the young person and DWP for the parents although this would require permissions from the relevant data holders and clear proposals of what data would be linked and why. So far, there are a few questions asked about expectations for the child's future plans at age 16 but no questions asking about career guidance/interventions. Wards with high levels of deprivation or high levels of ethnic diversity have been oversampled so this dataset could be especially useful for exploring these groups.
National Pupil Database (NPD)	Relevant topics include: pupil test and exam results, prior attainment and progression at different stages (early years, and Key Stages 1 to 5); information on pupils' characteristics including gender, ethnicity, first language, eligibility for free school meals, awarding of bursary funding, whether SEN, absence and exclusions. It does not capture attitudes or satisfaction metrics, and has only limited data on interventions (eg PLASC looks to contain a variable which identified whether there has been any planned employability, enrichment and pastoral hours in the current year [Non_qualHrs, collected from 2014/15 onwards] and in the previous academic year [Non_qualHrsPrev collected in 2014/15 only]).	This is a key useful dataset. It is however limited to England and does not include the other UK nations. It enables tracking of young people from a very early age to age 18 and by linking to ILR and HESA data tracks them beyond 18 so captures educational outcomes and short-term progression. It may be possible to link NPD with the new LEO dataset and to track individuals for a longer period of time to explore longer term careers.It does not capture attitudes or satisfaction metrics, and has only limited data on interventions (eg PLASC looks to contain a variable which identified whether there has been any planned employability, enrichment and pastoral hours in the current year [Non_qualHrs, collected from 2014/15 onwards] and in the previous academic year [Non_qualHrsPrev collected in 2014/15 only]). It is a highly complex dataset will require some guidance to use and to ensure the correct data extractions are requested

Data source	Relevance	Limitations of data source
OECD	OECD Employment and Labour Market statistics include several relevant datasets: labour force statistics, Earnings, Hours worked, Labour Market statistics, Job quality, Labour market programmes, and Average annual wages. Labour market statistics include (all with breakdown by sex and age (5 year or 15-24)): employment by job tenure intervals/permanency of the job, labour force statistics by sex and and unemployment by duration. Job quality includes (by gender, age or education groups) aspects such as: physical risk, emotional, training, opportunities to progress, autonomy, work/life balance, and satisfaction. Labour market programmes: expenditure and participants dataset includes: start-up incentives. The OECD Employment Outlook 2016 includes analysis of qualification mismatch and skills use. Archives include 2016 and 2015. OECD Education Statistics include several relevant datasets: Education at a Glance and UOE Education Database. Education at a Glance includes: profile of graduates and entrants for ISCED levels 3 to 8, share of enrolment by type of institution, part-time/full-time study intensity, Share of enrolment by gender, programme orientation and mode of study, transition from school to work for age groups (15-19, 20-24, 25-29) ie education/employed/unemployed/inactive, and educational attainment by labour force status. Historical data from 1997 onwards is available for: learning outputs and destination outcomes for different generations, and proportion of youth that are NEET. The UOE Education Database includes: enrolment by age, field, and type of institution; entrants by age and field, graduates by age and field, new entrants by sex and age, students enroled by age and type of institution. Historical data from 1997 onwards and presented by ISCED level is available for: graduates (age, field), new entrants by sex and age, students enroled (age, type of institution).	Data is collected at a national and not individual level so it is not possible to follow individuals through their education to work transition. Data is UK level and not broken down into UK nation. Education data is presented by ISCED level so it is only possible to get an idea of education stage rather than specific qualifications, or a clear picture of vocational/academic qualifications. Age data is presented as groups, e.g. 15-19, so it is not possible to identify typical transition points, e.g. 18 for the UK. OECD data is, however, useful to compare at a national level levels/types of education and employment status/activities across different countries.
Opinions and Lifestyle Survey (OPN), formerly Office for National Statistics (ONS) Opinions Survey or ONS Omnibus Survey. It became the Opinions and Lifestyle Survey in April 2012 following the merger of some elements from the General Lifestyle Survey (GLF), which closed in January 2012.	Each monthly module includes: core demographic questions, non-core questions where the topic varies from month to month. Government organisations, academic institutions and charities can commission a module on the monthly Opinions and Lifestyle survey. Non-core topics in 2016 have included well-being and forms of personal identification documents held. Core questions asked each time include: region, household, age left age when left full-time education, highest level of qualification, national identity, ethnicity, physical or mental health conditions, participating in a government scheme, paid work in the last 7 days, ever had paid work, managerial duties, employer size, Full-time/Part-time work, NS-SEC occupation info, income.	Participants in this study are adults aged 16+, no information is collected regarding young people in the household beyond the number living there. The topic of the survey changes with each wave but core information regarding current employment, highest qualification, income and personal characteristics are collected. However, the sample is randomly selected each time within set parameters so that the research does not track particular individuals over time. It could be used to ask adult populations about attitudes to or experiences of careers guidance/information, and if this were done in different years could be used to track trends in the general population over time.
School Performance Tables	Relevant topics include: attainment, assessment, progression/distance travelled (KS2, KS4, and ages 16-18), and destinations (KS4 and ages 16-18). Destinations data for 16 to 18 olds is a new development available from 2016. Also pupil numbers and profile, absence information, staffing numbers and profile and spend per pupil. There are a large number of documents to support usage (to interpret measures reported and to understand how data is captured and how this has changed over the years).	This is a set of tables that allows for some degree of interogation but is limited in the degree of manipulation permitted. Results for Individual schools and colleges, and individual local authorities can be produced and compared, and some degree of filtering is permitted by type of pupil (gender, level of achievement, disadvantage). Data can be downloaded which allows for a greater degree of manipulation but the data is still at an aggregate institution or local authority level. The data is not really suitable for linkage to individual level data. It covers England only.
Universities and College Admissions Service (UCAS)	Relevant topics include: age; ethnicity; school type; gender; socio-economic status; predicted A-levels; Application, course and HEP; provider decision; applicant response; unique RPD identifier.	UCAS holds data on all applicants to undergraduate HE courses and records quite a lot of useful information for understanding the trajectories of those entering HE. UCAS publish some aggregated data tables for a limited number of topics which are publically available. Access to further detailed data requires a bespoke data order for which a fee will be charged. The publically available data tables are of limited use to inform CEC's work; and more relevant variables and insight would be available through accessing a fuller dataset or bespoke data tables.

Data source	Relevance	Limitations of data source
UCAS Strobe	Relevant topics include: age; ethnicity; school type; gender; socio-economic status; predicted A-levels; Application, course and HEP; provider decision; applicant response; unique RPD identifier, which can be linked to the STROBE output via the UCAS Exact service, which again is a paid-for service. The key advantage of STROBE is that it allows tracking of specific individuals through the university application process, although results are presented in aggregate rather than at an individual level.	Cost of access is the major limitation, although it is thought that free access will become available via the Administrative Data Research Network. Also the data only cover the process of applying to university and accepting offers, and data on the outcomes ie which university/course attended, are available via NPD linked to HESA data.
Young People's Social Attitudes (YPSA)	Topics in the 2003 YPSA include: gender differences, problems at school, views about education and work, politics and decision-making, prejudice and morality, fulfilment, friends and social networks, and household tasks. Possible relevant questions include: volueteering/charity work, plans for moving out of parents' home, employment status, perceived gender difference relating to work and educational performance/interests, their attitudes to work, perceptions of important factors in being successful in life, main ambitions, classification/demographic info: employment status, nationality, ethnicity, religion, education, qualifications, educational institution,	The questionnaire provides three snapshots of attitudes among young people in Great Britain since 1998. It does not follow a cohort but collects a new sample each time. It does not capture any information about transitions, nor does it collect information about careers advice/guidance or interventions, but it does collected some information about attitudes to work and some future plans/aspirations.
Youth Cohort Study (YCS)	Relevant topics include: experiences of year 11, careers advice and guidance, employment status, education, NEET, background info on their family, attitudes, children or caring responsibilities, qualifications training for, sources of information, apprenticeships	The data presents a rich source of information regarding progression into education or employment in the four years following year 11 (age 16). However, the data cannot be used to explore long term outcomes so for individuals entering longer education programmes or higher education, the data does not capture the transition from further/higher education into subsequent employment or other outcomes. From 2007, the data can be linked to LYSPE data but prior to that this is not possible. Cohort thirteen can be also linked to the NPD. Prior to the 2005/6 data was collected for England and Wales, but is limited to England only from this cohort onwards.

Data source	Linking	Critical assessment of usefulness to CEC	URL?	Sample size?
Annual Population Survey (APS)	LFS/APS data has been combined with related data from other sources to provide an overall view of the state of the labour market. The Bank of England's Monetary Policy Committee, HM Treasury and the Department for Work and Pensions use LFS data in this way. Personal identifiers such as National Insurance number are not collected in the research but personally identifying information such as name, address, date of birth and sex are collected as part of the survey. The Postcode Address File is used for sampling so postcode is known prior to survey data collection.	Amber/Red	https://discover.ukdataservice.ac.uk/series/?sn=200002	c. 320,000 individual records per dataset
Annual Survey of Hours and Earnings (ASHE); formerly New Earnings Survey (NES).	NI number is collected for the employee so it is possible to link with national datasets to access more information about the employee. These data contain Inter-Departmental Business Register (IDBR) reference which makes it possible to link ASHE data with data from other business surveys, such as the Annual Business Survey (UK Data Archive SN 7451).	Amber	https://discover.ukdataservice.ac.uk/catalogue/?sn=6689&type=Data%20catalogue	Approximately 140,000-185,000 individuals per year
Aspires/Aspires 2	Part of the survey has been linked to the National Pupil Database (NPD) and the possibility of linking to HESA data is currently being explored. The linked NPD data will not be included with the data deposited with the UK Data Service once the research has been completed.	Green	http://www.ucl.ac.uk/ioe/departments-centres/departments/education-practice-and-society/aspires	Aspires: 4,600 to 9,319 young people, 65 parents. Aspires 2: 7,000 to 13,455 young people, 66 parents.
Avon Longitudinal Study of Parents and Children (ALSPAC), Children of the 90s; ALSPAC-G2, Children of the Children of the 90s, COCO90s	There is a dedicated team working on linking ALSPAC data with other data sources through the Project to Enhance ALSPAC through Record Linkage (PEARL). ALSPAC data has already been linked with health and environment data, and with educational datasets including: the National Pupil Database (NPD), the Pupil Level Annual School Census (PLASC), and Annual School Census (ASC). ALSPAC dataset has also been linked to HESA data and negotiations are currently in progress as to how this data can be made available to the wider research community.	Amber	http://www.bristol.ac.uk/alspac/	14,500 families in the Bristol area, 15,458 children from the age of 7.
British Cohort Study 1970, BCS70	There have been issues with consistency in identifiers used throughout waves of the data as different groups carried out the data collection at different stages of the study. However, there is documentation to clarify this though. A dataset has been developed for data collected 1970-2008 where occupation SOC codes have been added. Earlier waves of the study, e.g. around birth, appear to include clinical records which presumably includes patient data held by NHS. Some later waves included consent to link data to DWP, HMRC or NHS records.	Amber	https://discover.ukdataservice.ac.uk/catalogue/?sn=5641&type=Data%20catalogue	19,101 cases

Data source	Linking	Critical assessment of usefulness to CEC	URL?	Sample size?
British Council, Next Generation	The data does not collect any personal identifiers such as national insurance number or student numbers. Respondents are asked which university they attend so location of study can be inferred to some extent although it is possible that some may be participating in distance learning or campuses located elsewhere.	Red/Amber	https://www.britishcouncil.org/society/governance-civil-society/next-generation	2011 survey sample: 1,000, focus group sizes not given.
British Household Panel Survey (BHPS), 1991-2009, Understanding Society, United Kingdom Household Longitudinal Study; (UKHLS), 2009-2015	Adults in the Understanding Society (UKHLS) are asked about consent to link their data to DfE data, Health records NHS. The BHPS cohort is included in the UKHLS.	Green	https://discover.ukdataservice.ac.uk/catalogue/?sn=5151&type=Data%20catalogue	BHPS: 5,000 households UKHLS: Over 40,000 households
British Social Attitudes Survey (BSA)	The data does not collect any personal identifiers such as national insurance number or student numbers. As individuals are sampled on the basis of household it is assumed that postcode data is used for the sampling and some derived variables but the location variable provided is region.	Red/Amber	https://discover.ukdataservice.ac.uk/catalogue/?sn=8116&type=Data%20catalogue	2015 sample: 4,328 cases
Business and Community Interaction survey (HE-BCI)	HE-BCI data is collected at Higher Education Provider (HEP) level via HESA. It is therefore theoretically possible to link HE-CBI information about institutions to other HESA datasets which record HEP such as the student record.	Red/Amber	https://www.hesa.ac.uk/collection/c15032	All publically funded HEPs in UK, and participating alternative providers.
Census: Population census, and Microdata	Aggregated data does not contain individual level information and so cannot be linked. The Microdata does not include personal identifiers such as name, address, and date of birth so cannot be linked to other data. However, the original census data does contain personal identifiers such as name, postcode, date of birth and gender which could be used to link to some adminitrative data.	Amber/Red	https://www.ons.gov.uk/census/2011census	Entire population

Data source	Linking	Critical assessment of usefulness to CEC	URL?	Sample size?
Census: Neighbourhood statistics	The Neighbourhood Statistics link Census data with national administrative or survey data held by government departments. It has also been matched with the Millennium Cohort Study (MCS).	Amber	http://www.neighbourhood.statistics.gov.uk/dissemination/	N/A
Census: ONS Longitudinal Study	The Census Longitudinal Study for England and Wales links data for individuals and households across census waves from 1971 onwards. In England and Wales the longitudinal census data has been linked to data on registration of certain life events. These are: registration of birth, birth of child to mother, widowhood, cancer diagnosis, and registration of death. The original census data includes information on some personal identifiers such as postcode, date of birth, name and gender which could be used to link to some administrative data. Data linking is possible - researchers need to discuss proposed linking with CeLSIUS and they will decide whether it is appropriate and feasible.	Amber/Red		The LS represents a continuous sample of the population of England and Wales, rather than a sample taken at one time point only. It now includes records for over 950,000 study members.
Citizenship Education Longitudinal Survey (CELS) ; also incorporates Citizens in Transition: Civic Engagement among Young People, CiT;	The data has not yet been linked with orther datasets but the researchers are exploring the possibility of linking with the National Pupil Database (NPD).	Amber	http://www.llakes.ac.uk/research-project/166/citizenship-education-longitudinal-survey-cels	Circa 11,250 students in 2010.
Effective Pre-School, Primary and Secondary Education project (EPPSE)	The data incorporates information from the National Pupil Database (NPD), and PLASC.	Green	https://www.ucl.ac.uk/ioe/research/featured-research/effective-pre-school-primary-secondary-education-project	3,172 children and families across England, recruited in the pre-school period.
Employer Perspectives Survey	This dataset can be linked to the Employer Skills Survey.	Amber	https://www.gov.uk/government/publications/employer-perspectives-survey-2014	Around 18,000 employers

Data source	Linking	Critical assessment of usefulness to CEC	URL?	Sample size?
ESPON Database	As the data is presented at the national level, it is not possible to identify and track individuals. However, national data for the UK is compared with equivalent data from other countries - mostly EU but also some outside the EU such as the USA.	Amber/Red	https://www.espon.eu/main/	Population data.
European Longitudinal Study of Parents and Children (ELSPAC)	The UK dataset (ALSPAC) has been linked with national datasets such as the NPD. The Isle of Man ELSPAC data has not been linked to other datasets. The linking of ELSPAC datasets to relevant national data sources is decided and conducted at an inidvidual country level.	Amber/Red	http://www.cls.ioe.ac.uk/page.aspx?&sitesectionid=328&sitesectiontitle=European+Longitudinal+Study+of+Pregnancy+and+Childhood	Project ELSPAC comprises over 40,000 children and their families across the participating countries.
eurostat	Microdata containing individual level data is available and so tracking individuals over time through transitions may be possible if indivdiuals are given unique personal identifiers. National data for the UK is compared with equivalent data from other countries - mostly EU but also some others.	Amber/Green	http://ec.europa.eu/eurostat	Population data.
Families and Children Study (FACS); formerly known as the Survey of Low Income Families (SOLIF)	The study was commissioned by the Department for Work and Pensions (DWP) and the sample was recruited from individuals receiving Child Benefit. The research monitors income and benefits received and HMRC have compared some of this data with their administrative records. Permission has been sought in some waves to link data with DWP and NPD data.	Amber	https://discover.ukdataservice.ac.uk/catalogue/?sn=4427&type=Data%20catalogue	4,659 - 8,057 cases

Data source	Linking	Critical assessment of usefulness to CEC	URL?	Sample size?
Futuretrack Study	As part of the consent agreement to take part in the FutureTrack online survey, participants consented to their UCAS application data being transferred to the researchers in order to to be linked to their survey responses. The UCAS Personal Identifier can be used to link UCAS data with HESA data in order to track individuals through higher education.	Green		130,000 - 138,000
HESA Student Record	UCAS Data for HESE (previously known as J*) can be used to link HESA student records to UCAS applications. This also includes deferrals.	Amber (Green when linked to DLHE)	https://www.hesa.ac.uk/collection/c16051/	
HESA Destinations of Leavers from Higher Education (DLHE) survey, initial or early DLHE	DLHE data can and is routinely be linked to the HESA Student Record dataset using the 'HESA unique student identifier variable'. This enables background information such personal characteristics to be incoporated with information regarding the transition from higher education to employment or further study. For the NewDLHE there are plans to link this to salary data from the LEO dataset (see separate note), and to explore other data linkage opportunities. To enable linkages to LEO, HESA propose to hold and provide access to parts of the LEO dataset relating to earnings before, during and after graduation. It is hoped that the first LEO data linked into the NewDLHE will be available in June 2020.	Green	https://www.hesa.ac.uk/collection/c15018/	Most recent cohort (2014/15): 399,345 (76% response rate)

Data source	Linking	Critical assessment of usefulness to CEC	URL?	Sample size?
HESA Destinations of Leavers from Higher Education (DLHE) Longitudinal survey	DLHE data can and is routinely linked to the HESA Student Record dataset using the 'HESA unique student identifier variable'. This enables background information such personal and HE study characteristics to be incoprorated with information regarding the longer-term outcomes from higher education. LDLHE data can also be linked with the early DLHE results. NewDLHE: For the NewDLHE there are plans to link this to salary data from the LEO dataset (see separate note), and to explore other data linkage opportunities. To enable linkages to LEO, HESA propose to hold and provide access to parts of the LEO dataset relating to earnings before, during and after graduation. It is hoped that the first LEO data linked into the NewDLHE will be available in June 2020.	Green	https://www.hesa.ac.uk/collection/c10019/	Cohort 2010/11: 81,650 respondents (52,775 sample A, and 28,875 sample B). NB the sampled population to be surveyed from the 2010/11 cohort represented a larger sample than in previous surveys due to new criteria being introduced such as oversample of graduates funded by Health Education England, leavers with a degree in computer science, and leavers from disadvantaged areas measured using POLAR3 quintiles (12,930 compared with 80,835 for the 2008/09 cohort).
Individualised Learner Record, ILR	Learners are assumed to opt in, but have the option to opt out if they request. The ILR contains several fields that could enable linkage to other datasets: Name, Address, National Insurance Number and the Unique Learner Number. In 2010/11, the use of a Unique Learner Number (ULN) became a mandatory field in the ILR.The ULN is also collected as part of the Department for Education's Pupil Census, the Universities & Colleges Admission Service dataset and the Higher Education Statistics Agency Student Record dataset. More specifically, researchers can now also make a single application for linked data extracts from the National Pupil Database, HESA Student Records and Individualised Learner Records. There is also postcode before enrolment and during which could be linked to socio-economic data such as NS-SEC, ACORN, POLAR, etc. A 2011 study published by BIS linked ILR data with HMRC and DWP data: Patrignani, Pietro and Conlon, Gavan, Department for Business Innovation and Skills (BIS), corp creator. (2011) The long term effect of vocational qualifications on labour market outcomes. [Research paper number 47]	Green	https://discover.ukdataservice.ac.uk/catalogue/?sn=888021&type=Data%20catalogue	Approximately 12,500,000 records per year covering all further education funding streams
Labour Force Survey (LFS) Longitudinal Datasets	LFS data has been combined with related data from other sources to provide an overall view of the state of the labour market. The Bank of England's Monetary Policy Committee, HM Treasury and the Department for Work and Pensions use LFS data in this way. Personal identifiers such as National Insurance number are not collected as the research but personally identifying information such as name, address, date of birth and sex are collected as part of the survey. The Postcode Address File is used for sampling so postcode is known prior to survey data collection.	Amber	https://discover.ukdataservice.ac.uk/series/?sn=2000026	2Q dataset is c. 32,000 individuals; 5Q dataset is c. 5,000 individuals

Data source	Linking	Critical assessment of usefulness to CEC	URL?	Sample size?
Life Opportunities Survey (LOS)	Participants were asked for their consent to link their responses with information held about them by DWP. Data is linked using name, address, sex and date of birth, rather than National Insurance number. The survey does not appear to ask for any other personal identifiers so it may not be possible to link this data with other sources, and participants may need to be contacted to request permission for this.	Amber	https://discover.ukdataservice.ac.uk/series/?sn=2000027	48,135 cases across Waves 1-3.
Local authority interactive tool (LAIT)	LAIT is an interactive spreadsheet linking data from several Government departments and sources including: Local Authorities, Ofsted, DfE, Department of Health. In addition to this, the National Foundation for Educational Research (NFER) was commissioned in 2007 by the Department to identify and group similar LAs in terms of the socio-economic characteristics, each LA was assigned 10 'statistically similar' neighbours. However, it provides information at an aggregated level and not individual level (local authority) so it is not possible to identify and track individuals over time.	Amber	https://www.gov.uk/government/publications/local-authority-interactive-tool-lait	All children in England.
Longitudinal Education Outcomes (LEO)	The LEO datset links information from national datasets including: The National Pupil Database (NPD) held by the Department for Education (DfE); Higher Education Statistics Agency (HESA) data on students at UK publicly funded higher education institutions and some Alternative Providers held by DfE; Individualised Learner Record Data (ILR) on students at further education institutions held by DfE; Employment data (P45 and P14) held by Her Majesty's Revenue and Customs (HMRC); and The National Benefit Database, Labour Market System and Juvos data held by the Department for Work and Pensions (DWP). There are plans to link with the Work and Pensions Longitudinal Study (WPLS) and with the new DLHE survey, and work is underway to explore the potential to match with data regarding Universal Credit.	Green	https://www.gov.uk/government/statistics/graduate-outcomes-by-degree-subject-and-university	Graduates: Circa 200,000 to 278,000 per annual cohort. Key stage 5: 349,890

Data source	Linking	Critical assessment of usefulness to CEC	URL?	Sample size?
First Longitudinal Study of Young People in England (LSYPE1); Next Steps.	The LYSPE1 data has been linked with a number of sources for some variables: National Pupil Database (NPD), Index of Multiple Deprivation (IMD), Income Deprivation Affecting Children Index (IDACI), urban/rural indicators, School Level Data at baseline an KS2, National Statistics Postcode Directory (NSPD). Not all of this information is available in the dataset deposited with UK Data Service. Researchers requiring access to the fuller linked administrative files should contact DfE directly. Data are only available to approved researchers who make a request to DfE using the Confidentiality Agreement form available with LSYPE documentation on the UK Data Archive (http://www.esds.ac.uk/findingData/snDescription.asp?sn=5545#doc) or from team.longitudinal@education.gsi.gov.uk. later waves of LSYPE have included consent questions to enable linking between LSYPE data and data held by the Department for Work and Pensions and the Department for Business Innovation and Skills. Whilst consent has been obtained for this additional linking and was already held for additional attainment information beyond Key Stage 4, as of November 2011 these data are not yet available.	Green	https://discover.ukdataservice.ac.uk/catalogue/?sn=5545&type=Data%20catalogue	15,770 households at Wave 1, reducing to 8,682 households by Wave 7.
Second Longitudinal Study of Young People in England (LSYPE2); Our Future.	The LYSPE2 data has been linked with the National Pupil Database (NPD). NPD data is only included where pupils and parents have given full consent.There are three types of identifiers collected for schools which could theoretically be used to link to other datasets: primary sampling unit codes (PSUs), LAESTAB codes and Unique Reference Numbers (URNs).	Green (Future)	https://discover.ukdataservice.ac.uk/catalogue/?sn=7810&type=Data%20catalogue	13,000 young people
Millennium Cohort Study (MCS); Child of the New Century	Consent was sought from parents in Wave 5 to link data to DWP records. Individual pupil information for the Cohort Member has been linked for those who consented in England to the National Pupil Database and to similar data in Wales and Scotland - this is only available through restricted secure access.	Green (Future)	https://discover.ukdataservice.ac.uk/catalogue/?sn=7464&type=Data%20catalogue	13,287 parents; 7,430 teachers
National Pupil Database (NPD)	NPD is easily and routinely linked with Individualised Learner Record data and Higher Education Statistic Agency (HESA) student data and destinations data using the anonymised Pupil Matching Reference (aPMR) to track progression and outcomes. It can also be linked to UCAS data (to identify deferred entry to HE), National Client Caseload Information System (NCCIS) which records the activity of students including employment, training and NEET status, and the new Longitudinal Educational Outcomes (LEO) data to gather further details on employment and benefit status. There are a number of individual unique identifiers, but the key variable appears to be the anonymised Pupil Matching Reference (aPMR). As individual student identifying information can be extracted from NPD (name, address, date of birth etc, a Tier 1 request) it is possible to match bespoke data onto NPD such as Management Information on interventions, survey responses and particiation in Randomised Controlled Trials (experiments). This create a powerful tool to identify the impact of certain interventions	Green	https://www.gov.uk/guidance/national-pupil-database-apply-for-a-data-extract. Additional useful site: http://nationalpupildatabase.wikispaces.com/	Census not sample

Data source	Linking	Critical assessment of usefulness to CEC	URL?	Sample size?
OECD	Individual level data is not provided so it is not possible to link data from the education or employment datasets to other national level datasets, e.g. National Pupil Database.	Amber/Red	https://data.oecd.org/	Population data.
Opinions and Lifestyle Survey (OPN), formerly Office for National Statistics (ONS) Opinions Survey or ONS Omnibus Survey. It became the Opinions and Lifestyle Survey in April 2012 following the merger of some elements from the General Lifestyle Survey (GLF), which closed in January 2012.	The sample is drawn using NS-SEC information based on postcode. Particpants are not asked for personal identifiers such as NI number, education or health identifiers so indivdiual data cannot be linked to other national datasets such as those form DWP, DfE or NHS.	Red/Amber	https://discover.ukdataservice.ac.uk/series/?sn=2000043	Circa 900-1,000 respondents.
School Performance Tables	As an aggregate set of tables this does not appear to facilitate linking to other data sources	Amber	https://www.compare-school-performance.service.gov.uk/	Census not sample
Universities and College Admissions Service (UCAS)	UCAS Identifier is included in the HESA student record and may be used to track students who have entered higher education. There is potential to link data to other adminsitrative datasets through the Administrative Data Research Network.	Amber	https://www.ucas.com/corporate/data-and-analysis/ucas-undergraduate-releases/ucas-undergraduate-analysis-reports	Around 700,000

Data source	Linking	Critical assessment of usefulness to CEC	URL?	Sample size?
UCAS Strobe	There is the potential to link to other data via the Administrative Data Research Network although there are no details at present.	Amber	https://www.ucas.com/corporate/data-and-analysis/analytical-data-services/strobe	Entire application cohort each year - c. 500k UK domiciled applicants and c. 100k overseas applicants
Young People's Social Attitudes (YPSA)	The data does not collect any personal identifiers such as national insurance number or student numbers. As individuals are sampled on the basis of household it is assumed that postcode data is used for the sampling and some derived variables but the location variable provided is region. The YPSA is carried out in households where an adult has been interviewed for the BSA and was created to compare views of the young person and the adult in the household. It should therefore be possible to link the young person's data with that of adults in the same household.	Amber	https://discover.ukdataservice.ac.uk/catalogue/?sn=5250&type=Data%20catalogue	2003 sample: 997 (target) 663 (obtained)
Youth Cohort Study (YCS)	The methodology of the YCS changed at Cohort Thirteen, to allow linkage with the Longitudinal Study of Young People in England (LSYPE) (held at the UK Data Archive under SN 5545). Samples for Cohort Thirteen and LSYPE were taken from the same academic cohort and their questionnaires and methodology were harmonised to facilitate analysis of the merged datasets. The YCS data has been linked to administrative data held within the Department on the National Pupil Database (NPD).The available data include pupils' Key Stage 2, 3, and 4 attainment and data about the pupil such as free school meal eligibility and Special Education Need status. These data can be linked to the deposited survey data using a unique survey identification code. Researchers requiring access to the linked administrative files should contact the Longitudinal Surveys Team at the Department directly	Green	https://discover.ukdataservice.ac.uk/catalogue/?sn=6024&type=Data%20catalogue	Cohort 13 contains 7,525-4,829