



Department  
for Education

# **The National Pupil Database**

**Data Request Application Form**

# Introduction

## National pupil database data

This application form should be used for requesting extracts of data from the National Pupil Database (NPD). Other documents relating to the data request process can be found on the [NPD Online](#) website.

Before making an application, please make sure that you have read all the accompanying information on the [NPD Online](#) website including the [NPD User Guide](#).

This form should be filled in electronically and e-mailed to the NPD Data Request team at [NPD.REQUESTS@education.gsi.gov.uk](mailto:NPD.REQUESTS@education.gsi.gov.uk) with a completed [Information Security Questionnaire](#).

## Who can request access to the data?

The Education (Individual Pupil Information) (Prescribed Persons) (England) Regulations 2009, as amended enable the Department for Education to share individual pupil information from the NPD with named bodies and persons who for the purpose of promoting the education or well-being of children in England are conducting research or analysis, producing statistics, or providing information, advice or guidance. Access is subject to requesters complying with terms and conditions imposed under contractual arrangements and a rigorous approvals process.

Any person requesting access to the data will have to demonstrate that they will comply with all relevant requirements of the Data Protection Act 1998. They will need to demonstrate that they:

- Have appropriate security arrangements in place to process the data;
- Intend to use the data only for the specified purpose;
- Will keep the data only for the specified length of time;
- Will not try to further disclose the data without the prior written approval of the Department for Education.

Applicants will be expected to sign an agreement to confirm this and provide additional information about their security arrangements.

Please keep a copy of all the documents, including completed forms, for your own records.

## Linked data

If you are applying for linked data you need to apply using a different form which can be found online [here](#). Linked data is any combination of the following:

- NPD data linked to Individualised Learner Record (ILR) data;
- NPD data linked to Higher Education Statistics Agency (HESA) data;
- NPD data linked to ILR and HESA data; or
- ILR data linked to HESA data.

## Processing of NPD data

When you make a request for NPD data it will be considered for approval by the Education Data Division (EDD) with the exception of tier 1 data requests, which will be assessed by the department's Data Management Advisory Panel. The EDD will inform you of the outcome of the decision.

Where it is agreed to provide you with data, you will be provided with a [Data Sharing Agreement](#) and an [Individual Declaration Form](#) to sign and return.

## Publication of details of requests

Please note that the Department for Education (DfE) has a policy of publishing details of all requests it receives for NPD data on their website.

DfE will publish a description of the NPD data requested and whether or not it is requested to be linked to ILR data or HESA data, the name of the requesting organisation, the intended use of the data and details of the outcome. For the avoidance of doubt, no personal data will be published by DfE.

# Application Form

<b>1</b>	<b>Contact details</b>
<b>Name of Applicant</b>	
[REDACTED]	
<b>Name of person / organisation who will be entering into the Data Sharing Agreement ("the Requester")</b>	
UCAS	
<b>Organisation type (e.g. Academic institution, Research organisation, Media, Individual citizen, etc.)</b>	
Educational Charity	
<b>Position</b>	
[REDACTED]	
<b>Department</b>	
Policy and Research	
<b>Correspondence Address</b>	
Rosehill New Barn Lane Cheltenham Gloucestershire	
<b>Postcode</b>	
GL51 6EL	
<b>Telephone</b>	
[REDACTED]	
<b>Address at which the data will be processed (if different from above)</b>	
N/A	
<b>Postcode</b>	
N/A	
<b>Telephone</b>	
N/A	
<b>If you are a student please specify your supervising tutor and their telephone and e-mail address</b>	
N/A	

2 Details of DfE Project Sponsor (if applicable)	
Title (Mr, Mrs, Ms, etc.)	
Name	
Job Title	
Division / Team	
Contact Address	
Telephone	
E-mail address	

3 Have you (or your organisation) previously applied to use or access NPD data (Please insert X)	
Yes - state reference number(s), date of application(s) and name of applicant(s)	Reference numbers: supplies DR101217.01 [REDACTED] and DR121221.02 [REDACTED] [REDACTED]
No	

4 Information Commissioner's Office (ICO) Registration	
a	<b>ICO Registration Number or Exemption Ground</b> (It is mandatory to provide a registration number unless you have a valid exemption)
	Z7293924
b	<b>ICO Registration Expiry Date</b>
	3rd December 2014

5 Enquiry details	
NPD - Specify which datasets, versions, tiers, academic years and sensitive fields are required along with any specific requirements relating to data linking and coverage in line with the <u>NPD published guidance</u> .	

- Standard Tier 1 final KS5 extract 2011/12. To include all non-restricted spring PLASC/School Census and KS5 fields, along with KS5\_laestab, school postcode and school name, and the following identifying and sensitive data items (field names in parentheses)
  - Forenames (KS5\_forenames, forenames\_spr12)
  - Surname (KS5\_surname, surname\_spr12)
  - Middlenames (middlenames\_spr12)
  - Date of birth (KS5\_dob, dob\_spr12)
  - Output area (coa\_spr12)
  - Free school meals (fsmeligible\_spr12, everfsm\_all\_spr12)
  - Ethnicity (EthnicGroupMinor\_spr12, EthnicGroupMajor\_spr12, EthnicitySource\_spr12)
  - Child looked after (CLA\_31\_March\_spr12, CLA\_POC\_start)
  - Special educational need (SENprovision\_spr12, PrimarySENTtype\_spr12, SecondarySENTtype\_spr12 )
  - Unique learner number (ULN\_Census\_spr12)
  - Language group (LanguageGroupMajor\_spr12, LanguageGroupMinor\_spr12)
  - Anonymous pupil matching reference (KS5\_pupilMatchingRefAnonymous)

This standard anonymous extract should be matched to the spring PLASC/School Census standard Tier 4 extract, supplemented with the identifying and sensitive data items detailed above. No matches to prior attainment are required. Matching should be performed in such a way that records in the KS5 extract are not omitted if a link to the spring PLASC/Schools Census is not found (ie all KS5 standard extract records should be present). In cases where no match to the PLASC/school census are found the corresponding spring PLASC/school census fields should be blank.

- Standard final Tier 1 KS4 extract for 2011/12 matched to the spring PLASC/School Census. To include all non-restricted school census and KS4 fields, along with KS4\_laestab, school postcode and school name, and the following identifying and sensitive data items (field names in parentheses)
  - Forenames (KS4\_forenames, forenames\_spr12)
  - Surname (KS4\_surname, surname\_spr12)
  - Middlenames (middlenames\_spr12)

- Date of birth (KS4\_dob, dob\_spr12)
- Output area (coa\_spr12)
- Free school meals (KS4\_fsm, KS4\_fsm6, fsmeligible\_spr12, everfsm\_all\_spr12)
- Ethnicity (All ethnicity indicator fields, for example KS4\_wiri, KS4\_aind etc, EthnicGroupMinor\_spr12, EthnicGroupMajor\_spr12, EthnicitySource\_spr12)
- Child looked after (KS4\_cla\_pp\_6\_months, CLA\_31\_March\_spr12, CLA\_POC\_start)
- Special educational need (KS4\_senps, KS4\_sena, KS4\_sentype, SENprovision\_spr12, PrimarySENtype\_spr12, SecondarySENtype\_spr12 )
- Unique learner number (ULN\_Census\_spr12)
- Language group (KS4\_flang, LanguageGroupMajor\_spr12, LanguageGroupMinor\_spr12)
- Anonymous pupil matching reference (KS4\_pupilMatchingRefAnonymous)

This standard anonymous extract should be matched to the spring PLASC/School standard Tier 4 extract, supplemented with the restricted fields detailed above. No matches to prior attainment are required. Matching should be performed in such a way that records in the KS4 extract are not omitted if a link to the spring PLASC/Schools Census is not found (ie all KS4 standard extract records should be present). In cases where no match to the spring PLASC/School Census are found the corresponding spring PLASC/School Census fields should be blank.

- Standard final Tier 4 extract 2011/12 spring PLASC/School Census for all pupils in years 7 to 11 matched to any Key Stage 1-3 attainment data they may have. To include the anonymous pupil matching reference and the following Tier 1 and 2 fields
  - FSM eligibility (fsmeligible\_spr12, everfsm\_all\_spr12)
  - Date of birth (dob\_spr12)
  - Ethnicity (EthnicGroupMinor\_spr12, EthnicGroupMajor\_spr12, EthnicitySource\_spr12)
  - Language group (LanguageGroupMajor\_spr12, LanguageGroupMinor\_spr12)
  - 2001 Census Output Area (coa\_spr12)
  - Surname (surname\_spr12)
  - Forename (forename\_spr12)
  - Middle names (middlenames\_spr12)

- o Special Education Need (primarySENtype\_spr12, secondarySENtype\_spr12)
- o Child looked after (CLA\_31\_March\_spr12, CLA\_POC\_start)

In addition the school postcode, school name, school reference code, school type and school admission policy fields should be included, matched to any records of prior attainment (Key Stage 1-3, matched so that records without any prior attainment are included in the extract).

- Final Tier 4 extract 2011/12 Spring PLASC/School Census for all pupils, including details of the pupils school. Unmatched - to include PupilMatchingRefAnonymous.
- 2012 institution history - school level, unmatched - to include leaestab in order to link
- 2012 institution Link - school level, unmatched - to include leaestab in order to link
- Individual Key Stage 5 Results Table data for 2011/12, including all Tier 4 fields listed under the "KS5 Result Table" section in the NPD data tables spreadsheet (including KS5\_PupilMatchingRefAnonymous to enable linking to other extracts outlined above). Unmatched. There should be multiple rows per student, one for each qualification awarded, with an anonymous pupil matching reference. To be presented as a single .txt file. this request is the same as that made by HEFCE (DR130830.01).
- Individual KS4 Results Table data for 2011/12, including all Tier 4 fields listed under the "KS4 Result Table" section in the NPD data tables spreadsheet (including KS4\_PupilMatchingRefAnonymous to enable linking to other extracts outlined above). Unmatched. There should be multiple rows per student, one for each qualification awarded, with an anonymous pupil matching reference. To be presented as a single .txt file. this request is the same as that made by HEFCE (DR130923.02).

In addition to the extracts outlined above, we seek permission to extend our use of the extracts received under DR121221.02 and DR101217.01 for the same period as for the extracts requested above. The remainder of this form relates to the data previously requested under both of the references given above along with the request for new data extracts relating to the 2011/12 academic year.

6	Intended use of data
a	What are the aims of your project or research?

UCAS is the organisation responsible for managing applications to higher education courses in the UK. UCAS have committed to transform the information value of its data assets by analysing them in the context of other data sources. By linking UCAS data back to pre-higher education records, and forward to higher education and destination data, better intelligence for decision making will be provided to Government, university and individual audiences. Examples of the way in which NPD data has been used to enhance UCAS applications data can be found [here](#).

In broad terms the type of analysis that the NPD data will be used for is:  
 - Adding extra descriptor variables to applicants to enable analysis of the application process by these new dimensions;



- Using the NPD itself as a denominator, linked or otherwise, to estimate application process related outcomes against the NPD base;
- Using the attainment and demographic information embedded within the NPD, in conjunction with any analysis derived patterns, to develop intra and inter cycle extrapolation and forecasts;
- Using the information in the linked and unlinked NPD for the purposes of establishing individual and group level real and synthetic comparison groups for the assessment of actual and proposed interventions or changes to the application process.
- To use the NPD to form control groups for the purposes of helping to monitor impacts of widening participation activities and other policy relevant education initiatives.

UCAS have been asked to take forward this programme of analysis by the Minister of State for Universities and Science in his speech to the Universities UK conference on the 9 September 2010 (<http://www.bis.gov.uk/news/speeches/david-willetts-uuk-conference>).

This data and analysis framework is being developed in conjunction with the Office of Fair Access and BIS analysts so that it can be used to assess the impact of the new tuition fee and student support arrangements.

Our work will likely contribute to the work of the Product Development Unit at DfE ( [REDACTED], Product Development Unit, [REDACTED], email: [REDACTED] in looking at Destination Measures (for example see [here](#)). These aims will be achieved by the sharing of linked aggregate data with this unit.

There will be a range of outputs from the use of the NPD data in this analysis framework. These will range from national level research findings and statistics (for example, tracking HE application rates by NPD-sourced background measures, and measures of the success of HE related choices), through statistics and planning tools for universities to aid them in delivering Government objectives, to data-led information and guidance to young people and their schools during the UCAS application process.

National and institutional outputs would be designed to eliminate any disclosure risk, by meeting ONS disclosure standards. They will typically take the form of statistical indicators (for example, the average grades expected to be required for admission this cycle, or how many acceptances an institution's pattern of offers is expected to generate) or research insights (for example; Have variable fees influenced the choice of institution? How important is GCSE mathematics to gaining a place on a Physics HE course?) A typical published output would be a table in the UCAS end-of-cycle annual report, for example application rates of English school leavers by type of school.

These insights will also be used to help provide more relevant information to individuals to support their higher education related decision making. For instance, analysis of previous cohorts might suggest higher education choices that lead to positive outcomes for young people with particular profiles of GCSE results: this insight might be used to inform people whose profile of attainment recorded on the UCAS application form indicates they are similarly placed.

BIS have approached UCAS with a request for access to linked NPD-UCAS data. UCAS seek permission here to share such linked data with BIS, and also the Higher Education Funding Council for England in anticipation that a similar request will be made by them, for internal analysis and research purposes only. The data would be a list of anonymous UCAS individual identifiers relating to individuals which were linked into the NPD (using UCAS's linking methodology) along with their anonymous pupil matching reference (taken from the NPD). Should permission for this use not be granted by DfE then UCAS would not wish this to jeopardize the remainder of this request and would ask that approval could be considered for all other uses outlined here.

<b>b</b>	<b>Is there a specific question you are seeking to answer?</b>
----------	--

There are several broad questions that we hope to help answer. These include

- Assessing the impact of the new tuition fee and student support arrangements
- To assess the impacts of widening participation activities and other policy relevant education initiatives.
- Exploring with the DfE whether our experiences with linking the NPD to the UCAS data under this agreement might form the foundation for DfE to incorporate elements of UCAS data into the National Pupil Database framework for the benefit of a wider group of policy and academic researchers.
- Whether relevant statistical indicators can be developed to aid the decision making process of institutions and applicants.

**c** Who is the intended audience?

Multiple audiences:

- Nationally published research,
- UCAS member institutions,
- UCAS applicants and the centers through which they apply,
- External organisations such as Teach First

**d** Why do you need the data requested to complete the aims of your project or research? You must specify why it is necessary for you to have the data you have requested including if appropriate whether you are conducting research or analysis, producing statistics or providing information, advice or guidance and how this will work will promote the education or well-being of children in England.

The individual level data extracts we request are required in order to meet the aims of the work set out above. For example without detailed information about individuals' characteristics while at school it would be impossible to assess the impact of the new tuition fee and student support arrangements on different groups of applicants.

It is hoped that the research and analysis, statistics and information generated through this work will promote the education and/or well-being of children in England via enabling prospective HE students and HE institutions to make better choices during future application cycles, leading to improved outcomes for both.

**e** Where fields constituting sensitive or disclosive data are requested, please explain on a field by field basis why each item is required and why the same research outcome cannot be achieved by using less sensitive or disclosive data

Date of Birth, names, output area and Unique Learner Number are used to form analytically robust links into the UCAS applicant data.

Date of birth is required as one of the primary individual level keys needed to match into the UCAS data (the others being names, output area, sex and Individual Learner Number (where assigned)).

The replacement with a less sensitive field such as age, even if measured in months, seriously weakens the ability to match between the records in cases where there has been geographical movement prior to HE entry through UCAS (this occurs frequently).

Output area is required to form matches where name details are not unique or have altered (which is common). The best matching would be obtained through using postcode but we have calculated that using output area causes a relatively small reduction in quality of matching compared to that of postcode.

In addition the fine geographical resolution of output area allows better matching of area-based statistics - especially those based on Census geographies preceding or succeeding the 2001 Census. The importance of fine level area analysis of policy relevant analysis of HE entry is demonstrated by the HEFCE and OFFA reports in this area ([http://www.hefce.ac.uk/pubs/hefce/2010/10\\_03/](http://www.hefce.ac.uk/pubs/hefce/2010/10_03/), <http://www.offa.org.uk/wp-content/uploads/2010/11/Trends-in-young-participation-by-selectivity-of-institution.pdf>).

Ethnic group (mapped), language group (mapped), special education needs and looked after indicators are required to create measures (such as HE application rates) by these policy relevant indicators.

Special education needs are analysed as a proxy to certain types of disability (for a recent example of this approach, using NPD data linked to a portion of the UCAS data held by HEFCE (under request DR090701.02 and related) see [http://www.hefce.ac.uk/pubs/hefce/2010/10\\_32/10\\_32.pdf](http://www.hefce.ac.uk/pubs/hefce/2010/10_32/10_32.pdf)).

<b>f</b>	<b>Are you proposing to match the data you have requested with any other data?</b> (Please insert x)	
Yes - Please set out your intentions and provide copies of any approval(s) you have obtained	<b>X</b>	<b>We intend to match the data with internal HE applications data held on UCAS's systems</b>
No		
<b>g</b>	<b>If the data is to be published or reproduced (having applied agreed disclosure control measures as in section k below) what format will this be in?</b> (Please insert an x for all that apply)	
Academic research paper	<b>X</b>	
Internal publication or report	<b>X</b>	
Website general access	<b>X</b>	
Website with restricted access	<b>X</b>	
Commercial publication for which no charge will be made	<b>X</b>	
Commercial publication for which a charge will be made	<b>X</b>	
Other (please detail)		
<b>h</b>	<b>Details of other individuals who you propose to have access to the data including their name, job title,</b>	

organisation and address and why it is necessary for them to have access. Please note each user is required to make an individual declaration and may be required to complete a separate Information Security Questionnaire.

Name	Job Title	Reason for access to data
[REDACTED]	[REDACTED]	Analysing linked UCAS/NPD data and aggregated NPD data
[REDACTED]	[REDACTED]	Analysing linked UCAS/NPD data and aggregated NPD data
[REDACTED]	[REDACTED]	Analysing linked UCAS/NPD data and aggregated NPD data
[REDACTED]	[REDACTED]	Analysing linked UCAS/NPD data and aggregated NPD data
[REDACTED]	[REDACTED]	Analysing linked UCAS/NPD data and aggregated NPD data

**i** Where the individuals listed in "h" are from a different organisation to the applicant is there a contract in writing between the applicant and the other organisation(s) and does this cover data protection and information security arrangements?

<b>Yes</b> - Provide a summary of the contractual relationship and the data protection and information security arrangements in place.	
<b>No</b> - Explain why no written contract is required and provide us with a summary of the data protection and information security arrangements in place?	
<b>Not applicable</b> - No additional organisations are listed in "h".	

**j** When will you commence using the data? How long are you seeking to retain the data for? Why is this period necessary?

We will commence using the data immediately upon receipt. We seek to retain the data for as long as the analyses outlined here are deemed necessary by Government and others. This may be for several years given that many of the analyses outlined gain substantially in analytical and policy relevance when conducted for either a substantial time series, or when they span a substantial time gap, for example between intervention and outcome.

**k** Are you proposing to use the Standard Disclosure Control as set out at section 4.3 of the NPD User Guide? (Please insert X)

Yes	
No - How will you ensure that the confidentiality of individuals is preserved in any outputs arising from your use of the data?	Our disclosure control will meet ONS standards for disclosure control (see <a href="http://www.ons.gov.uk/ons/guide-method/best-practice/disclosure-control-policy-for-">http://www.ons.gov.uk/ons/guide-method/best-practice/disclosure-control-policy-for-</a>

7 Have you discussed your request in advance with the DfE? If so, please specify the members of staff who have assisted you.

This request has been discussed via email with [REDACTED] of the National Pupil Database & Transparency Unit. In response to this we have completed this form and the Information Security Questionnaire.

Also note that this request is very similar to our previous request DR121221.02. The main difference is that here we request data from the more recent 2011/12 academic year. For our previous request we corresponded with [REDACTED] in the Product Development Unit.



Department  
for Education

© Crown copyright 2013

You may re-use this document / publication (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence v2.0. To view this licence, visit [www.nationalarchives.gov.uk/doc/open-government-licence/version/2](http://www.nationalarchives.gov.uk/doc/open-government-licence/version/2) or email: [psi@nationalarchives.gsi.gov.uk](mailto:psi@nationalarchives.gsi.gov.uk).

Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

Any enquiries regarding this publication should be sent to us at [www.education.gov.uk/contactus](http://www.education.gov.uk/contactus).

This document is also available from our website at [NPD Online](#).