

UNCLASSIFIED



Department  
for Education

# **The National Pupil Database**

## **Data Request Application Form**

UNCLASSIFIED

<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>	
BBC NEWSNIGHT – PART OF THE BRITISH BROADCASTING CORPORATION	
<b>Organisation type (e.g. Academic institution, Research organisation, Media, Individual citizen, etc.)</b>	
MEDIA	
<b>Position</b>	
[REDACTED]	
<b>Department</b>	
BBC NEWSNIGHT	
<b>Correspondence Address</b>	
[REDACTED]	
BBC BROADCASTING HOUSE PORTLAND PLACE LONDON	
<b>Postcode</b>	
W1A 1AA	
<b>Telephone</b>	
[REDACTED]	
<b>Address at which the Linked Data will be processed (if different from above)</b>	
N/A	
<b>Postcode</b>	
<b>Telephone</b>	
[REDACTED]	
<b>If you are a student, please specify your supervising tutor and their telephone number and e-mail address</b>	

<b>2</b> Have you (or your organisation) previously applied to use or access NPD or linked data? (Please insert X if 'No')	
Yes - state reference number(s), date of application(s), and name of applicant(s)	YES, VARIOUS APPLICATIONS VIA THE FINANCIAL TIMES 2010-2012
No	
<b>3 Information Commissioner's Office (ICO) Registration</b>	
<b>a</b> ICO Registration Number or Exemption Ground (It is mandatory to provide a registration number unless you have a valid exemption)	
Z517352X	
<b>b</b> ICO Registration Expiry Date	
4 January 2015	
<b>4 Enquiry details</b>	
<b>a</b> 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> .	
<p>KS4 – Final – 2013, 2012, 2011, 2010, 2009 Matched to: School census, KS2, KS1 prior attainment</p> <p>KS2 – Final – 2013, 2012, 2011, 2010, 2009 Matched to: School census, KS1 prior attainment</p> <p>(10 datasets)</p>	
<b>b</b> ILR - Specify the academic years and individual fields required. The available linked data is listed <u>here</u> .	
<b>c</b> HESA - Specify the academic years and individual fields required. The available linked data is listed <u>here</u> .	

<b>5 Intended Use of Data</b>	
<b>a</b> What are the aims of your project or research?	
<p>To cover the English school system for BBC Newsnight, a BBC2 TV programme, and its associated online outlets. I would like to cover the progress made by the current - and previous - administration in reducing geographical, social and economic disparities in education performance.</p> <p>In my previous work, I have basically drawn up every conceivable indicator for 2007-12. I would like to be able to</p>	

update this. I would also like to be able to assess the structural reforms, and assess providers – such as chains.

**b** Is there a specific question you are seeking to answer?

I am particularly interested in geographical and social differences in school outcomes.

**c** Who is the intended audience?

BBC Newsnight's audience and the readership of BBC Online

**d** Why do you need the data requested to complete the aims of your project or research? If you are requesting NPD data 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 work will promote the education or well-being of children in England. For access to ILR and HESA linked data you will need to include in your explanation the relevant condition(s) under Schedule 2 and/or 3 of the Data Protection Act 1998, which you consider provides the justification for the data requested to be provided to you. Schedule 2 & 3 of the Data Protection Act can be found at: <http://www.legislation.gov.uk/ukpga/1998/29/schedule/2>  
<http://www.legislation.gov.uk/ukpga/1998/29/schedule/3>

I will be producing research and analysis on school performance, the impact of policy, ethnicity, poverty and neighbourhood – and how they interact. So school-level data is not helpful. School-level data also opens the possibility of ecological fallacies (ie – not knowing who is doing better or worse from certain effects).

I appreciate that this is a lot of data, but one key idea I would like to examine is the structural reforms. And, in simple terms, I need to know what direction local schools were moving in at the moment at which they converted.

To be more concrete, I am proposing quite an involved bit of modelling. The government believes competition can raise effects. That means the effect of academisation may occur in neighbouring schools, not just academies. So I wish to test for this by examining the effect of academy conversion at a neighbourhood level (an important reason why school-level data won't suffice). So I would like to build a model where the dependent variable is *neighbourhood* performance at the output area level.

For each neighbourhood, I could use the NPD to build an entry for the panel for each year. Then, alongside it, I could fill in data about the opening of new schools, academy conversion and so on in the local area. Then I intend to build a model which will incorporate this so that I can get a coefficient that says "if your nearest school converts, expect a [x] per cent rise in GCSE points."

I would cheerfully request less, but if I don't have the 2009 data, I can't make the panel incorporate the first conversions (which are a helpful discontinuity). And if I don't have data through to 2013, I won't have the benefit of giving the schools enough time to respond.

**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 using less

**sensitive or disclosive data.**

I would like the following sensitive fields (most of which are in tier 2 releases):

- Gender\_[term][yy]
- AgeAtStartOfAcademicYear\_[term][yy]
- MonthPartOfAgeAtStartOfAcademicYear\_[term][yy]
- Ethnicity\_[term][yy]
- EthnicGroup\_[term][yy]
- EthnicGroupMinor\_[term][yy]
- EthnicGroupMajor\_[term][yy]
- FSMeligible\_[term][yy]
- EVERFSM\_ALL\_[term][yy]
- COA\_[term][yy]
- LSOA01\_[term][yy]
- IDACIScore\_[term][yy]
- FirstLanguage\_[term][yy]
- EVERFSM\_6\_[term][yy]
- InCare\_[term][yy]

- ServiceChild\_[term][yy]
- GandTindicator\_[term][yy]
- ModeOfTravel\_[term][yy]
- Disability\_[term][yy]
- SENprovision\_[term][yy]
- Primary SEN
- DistCurrSch\_[term][yy]
- DistNearSch\_[term][yy]

The tier 1 variables I have requested are

- **Output area** - This is important for analysing the effect of policies such as academy conversion. Using output areas, I can get the child's neighbourhood with greater precision. When analysing competition effects, for example, knowing the nearest schools with greater precision is important. LSOAs are slightly too large for comfort, so the child may not be allocated to the right network of local schools for analysis.

- **Ethnicity and mother tongue** – this is an addition to the request: I would like these variables because, as [REDACTED] has argued this week, I don't think we quite understand the ethnic and cultural components of the London school improvement. I can't spot the effect of, say, A8 immigration on schools because I can't spot Poles, say, as opposed to Germans. Nor can I distinguish Somali and Nigerian children (who are rather different).

- **In care** – this is a crucial group, who suffer multiple disadvantages; assuming I can meet the anonymisation rules (which might mean not passing comment on lots of LAs to avoid breaching the small number rules), I would like to take a look at differentials in care. Given that some of these children physically get moved around the country a lot, this is particularly important (if we shipping children to schools with weak LAs.)

- **Disability** – This is because I think part of the London advantage is the manner in which SEN is endogenous to the schools. A more objective indicator might help pick that up, even if it isn't enough to regress on.

-Service children – They now attract extra money, so I would like to know if they have any recognisable differences. This just strikes me as being worth examination.

- Gifted and talented – These pupils have particular needs whose experiences, again, would bear examining.

- Distance from schools & mode of travel – School travel bears examining, especially since the DfE has sought to change traffic planning rules. I'd also like to know whether the decision to eschew nearer schools/drive to school is predictive of anything. (From the output areas, I will be able to ascertain the distances anyway. But since you have calculated them...!)

- Primary SEN If areas display dramatically different treatments for the same recorded disadvantage, that will help decisions about whether to permit SEN as a factor. In general, I try to avoid it because I think it is endogenous to the schools. If there's consistency, I might rethink that.

Among the other sensitive data, I am requesting the usual variables which can help to explain differences in school performance. I appreciate that this :

- Broad ethnic groups
- EAL status
- FSM eligibility
- SEN
- Age
- Gender

<b>f</b>	<b>Are you proposing to match or link the linked data you have requested with any other data?</b> (Please insert X if 'No')	
Yes - Please set out your intentions and provide copies of any approval(s) you have obtained?		
No		No
<b>g</b>	<b>If the data is to be published or reproduced (having applied agreed disclosure control measures see section k) what format will this be in?</b> (Please insert an X for all that apply)	
Academic Research Paper		
Internal publication or report		
Website General Access		X - BBC News
Website with Restricted Access		
Commercial publication for which no charge will be made		
Commercial publication for which a charge will be made		

Other (please detail) Television.

**h** Details of other individuals who you propose to have access to the data including; Name, Job Title, Organisation, 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	Organisation	Address	Purpose

**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".	N/A

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

I would like to hold it for six months. This would allow me time to perform the analyses I need to do, and to have it on hand to respond to policy announcements.

**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 if 'Yes')

<b>Yes</b>	YES
<b>No</b> - How will you ensure that the confidentiality of individuals is preserved in any outputs arising from your use of the linked data?	

**6** Have you discussed your request in advance with DfE, SFA, BIS, HESA or HEFCE if so please specify the members of staff who have assisted you.



**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 [NPD.REQUESTS@education.gsi.gov.uk](mailto:NPD.REQUESTS@education.gsi.gov.uk).

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