

Personal data: access to identifying and/or highly sensitive data

The Department for Education gives away the personal data of ca 20m people¹.

The table below is extracted from *the National Pupil Database User Guide*² [pages 19-20]. Only tier 3 data are aggregated but may still be identifying due to "small numbers and single counts." Tiers 1,2 and 4 are releases of individual identifiable pupil level data.

Tier 1: Individual pupil level data – identifiable and / or identifiable and highly sensitive

Individual pupil level extracts that include identifying and highly sensitive information about pupils and their characteristics including items described as 'sensitive personal data' within the Data Protection Act 1998.

Tier 2: Individual pupil level data – identifiable and sensitive

Individual pupil level extracts that include sensitive information about pupils and their characteristics including items described as 'sensitive personal data' within the Data Protection Act 1998 which have been recoded to become less sensitive.

Tier 3: Aggregate School level data – identifiable and sensitive

Aggregated extracts of school level data from the Department's school level database which could include items described as 'sensitive personal data' within the Data Protection Act 1998 and could include small numbers and single counts.

Tier 4: Individual pupil level data – identifiable

Individual pupil level extracts that do not contain information about pupils and their characteristics which is considered to be identifying or described as sensitive personal data within the Data Protection Act 1998.

The Department for Education third party release register³ lists data recipients 2012-14. These include releases of fully identifying data at individual level including name. These data are personal data. The DfE user guide states:

"where access to identifying and/or identifiable and highly sensitive data items is required, approval will be sought from the department's Data Management Advisory Panel (DMAP)." [see footnote 2: p26]

Of the 462 releases that have been through the DMAP process in 2012-2014, only 53 have been aggregated data:

163 Tier 1 - individual level, identifiable and highly sensitive (2 rejected)

177 Tier 2 - individual level, identifiable and sensitive (1 rejected)

11 Tier 3 - identifiable, sensitive and aggregated, may be identifying in small numbers

53 Tier 4 - individual level, identifiable, non-sensitive items (1 rejected)

60 unclassified (include sensitive and identifiable data) including 5 that were rejected.

¹ https://www.whatdotheyknow.com/request/pupil_data_national_pupil_databa_2#incoming-764676

² https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/472700/NPD_user_guide.pdf

³ <https://www.gov.uk/government/publications/national-pupil-database-requests-received>

Individual and Named Releases of Personal Data

Example 1: Use of name and direct intervention using the data

In February 2014, the IoE were given the names of year 7 pupils (age 11) in England for researchers to carry out a named maths survey [line 288, register]

*"Schools will be sent scripts with children's name on, complete the test and send it back to GL Assessment. **IOE require a complete list of pupil names of all year 7 pupils within the schools to facilitate this testing.** They are also intending to use the NPD data to measure prior attainment of children..."*

Example 2: Use of name and direct intervention using the data

NPD named personal data were extracted and matched with health data held by the HSCIC, for an individual named postal survey "What About YOUth" in 2014.⁴ An intensely detailed social survey mailshot was sent to the homes of nearly 300,000 15 year old pupils. [See the published report.⁵]

A random sample of all those pupils who turned 15 in academic year 2013/14, whose dates of birth ranged between 01/09/1998 and 31/08/1999, was selected. At the start of the trial survey, two potential sample frames were available from which to draw a sample of 15 year olds in England. These were MIDAS⁵ and the National Pupil Database (NPD)⁶. As permission to use MIDAS for the main stage survey was not granted, the sample was drawn from NPD, specifically the 2014 Spring School and Alternative Provision Censuses. Even though it excluded pupils in independent schools, the NPD was used as a sampling frame as it contains a broader range of matched background variables that could be used in the stratification of the sample. The decision to use NPD, without the inclusion of independent schools (just under 7% of the total number of school children in England), was taken after concluding that the estimates produced by the survey, would not be significantly impacted as a result of these schools not being included in the sample frame.

A sample of 298,080 young people was drawn from the NPD.

Example 3: Name, address and date of birth released to Cabinet Office

In June 2011, and 2013 the Cabinet Office extracted the names and addresses of 16-18 year olds, for piloting an approach looking at the completeness of the electoral register. [the latter is missing from release register but comment published via FOI, p8/36].⁶

Any other details? e.g. have you previously worked with any other Government Departments?
--

CO has previously accessed NPD data including name, address and date of birth of 16-18 year olds for the purposes of piloting an approach aimed at improving the completeness and accuracy of the electoral register. Data was shared as part of the "Cabinet Office Data Matching Pilot Schemes" firstly in June 2011 and subsequently in March 2013.
--

⁴ <http://www.whataboutyouth.com/who-we-are.aspx>

⁵ <http://www.hscic.gov.uk/article/3742/What-About-Youth-Study>

⁶ <https://www.whatdotheyknow.com/request/293030/response/723407/attach/6/Cabinet%20Office%20060913.pdf>

Personal data released at individual level include sensitive data

Items can be made 'less sensitive' but remain identifying, at individual level.

2	Individual pupil level data – Identifiable and Sensitive	<p>Individual pupil level extracts that include sensitive information about pupils and their characteristics including items described as 'sensitive personal data' within the UK Data Protection Act 1998 which have been recoded to become less sensitive.</p> <p>Examples of sensitive data items include ethnic group major, ethnic group minor, language group major, language group minor, Special Educational Needs and eligibility for Free School Meals.</p>	☒
---	--	---	---

According to the Department for Education own description there are some items which are released but are not possible to make less sensitive.

These include candidate numbers, names, ethnicity and disability, special needs, detailed breakdowns of special educational needs, whether the child has local authority looked after status, are children of military service personnel, free school meals, their reasons for absence or exclusion and whether they are recipients of free school meals.

Data items that cannot be made less sensitive.	Exam Candidate Number
	Candidate ID
	Names
	Primary and Secondary special educational need (SEN) type, e.g. specific learning difficulty, hearing impairment etc.
	Special Educational Needs
	Service children in education indicator
	Pupil's type of disability
	Children in Need
	Children Looked After (CLA)
	Pupil Premium
	Free School Meals (FSM)
	Reasons for Absence
	Reasons for Exclusions

Sensitive, Identifying Personal Data released to Journalists

Example 1: TV journalist, identifiable, highly sensitive personal data

A BBC Newsnight journalist in August 2014 was given Tier 1 identifying and highly sensitive data. They describe in their application how they will take small number rules into account using the data, because they are identifying and "School-level data is not helpful."

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).

Ethnicity was released at individual level to separate within ethnic groups:

- **Ethnicity and mother tongue** – this is an addition to the request: I would these variables because, as [REDACTED] [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).

Processing sensitive data requires additional DPA conditions to be met.⁷ The business cases in applications by journalists do not contain any indication how they meet schedule 3 conditions.

There is no proven 'need' nor 'benefit' to the individuals of releasing these identifying sensitive data like SEN or ethnicity. No audit had ever been done at the time we asked DfE this, to identify whether any benefits of the data release [see FOI 20 Aug 2015]⁸ were achieved to meet the legal requirement for release: "promoting the education or well-being of children in England."⁹

Is the justification data "bears examining" sufficient to release millions of children's identifiable, sensitive personal data?

*A full copy of the data application request was obtained through a Freedom of Information Request and is publicly available via What Do They Know.*¹⁰

-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

⁷ <https://ico.org.uk/for-organisations/guide-to-data-protection/conditions-for-processing/>

⁸ https://www.whatdotheyknow.com/request/pupil_data_application_approvals#outgoing-482241

⁹ <http://www.legislation.gov.uk/ukxi/2013/1193/regulation/2/made>

¹⁰ <https://www.whatdotheyknow.com/request/293030/response/723407/attach/10/BBC%20Newsnight.pdf>

Example 2: Newspaper journalist, identifiable, sensitive personal data

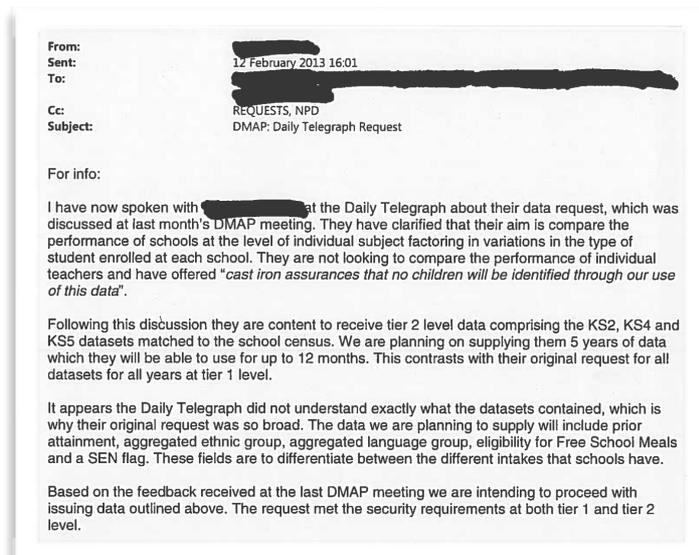
The Times journalists¹¹ were granted sensitive identifying, individual level data in September 2013. One of the purposes requested was, *"to pick interesting cases / groups of students."* [This is an extract from their data request form]

Please select the suitable tier which will cover all the information you are applying for

Information Tier	Information Type	Description	Tier Requested
1	Individual pupil level data - Identifying and / or Identifiable and Highly Sensitive	<p>Individual pupil level extracts that include identifying and highly sensitive information about pupils and their characteristics including items described as 'sensitive personal data' within the UK Data Protection Act 1998.</p> <p>Examples of identifying data items include Names, Address and Date of Birth.</p> <p>Examples of highly sensitive data items include Looked After Status, In Need Status, Full Ethnicity, Full Language and Primary and Secondary SEN Type, reasons for exclusions and absence.</p>	<input checked="" type="checkbox"/>

Example 3: Newspaper journalist, identifiable, sensitive personal data

Ten Telegraph journalists were given the personal data of ca. 10m children in February 2013.¹² The newspaper offered *"cast iron assurances that no pupil will be identified through our use of the data."* and received 5 years worth of identifying individual level and sensitive Tier 2 data. [see letter linked below ¹³] If the data were not identifying there would be no need to offer this assurance.



¹¹ <https://www.whatdotheyknow.com/request/293030/response/723407/attach/5/The%20Times.pdf>

¹² <https://www.whatdotheyknow.com/request/293030/response/723407/attach/3/Daily%20Telegraph.pdf>

¹³ <https://www.whatdotheyknow.com/request/293030/response/738135/attach/2/Annex.pdf>

Fair Processing of data released from the National Pupil Database:

The DfE privacy template: <https://www.gov.uk/government/publications/data-protection-and-privacy-privacy-notice> **links to this data usage statement:** <https://www.gov.uk/guidance/data-protection-how-we-collect-and-share-research-data>

There is no mention on the DfE privacy template for schools, the links from it, or their video to schools with instructions, that tells schools or parents that children's identifiable, individual data may be given to commercial third parties, TV or Fleet Street journalists, charities or data management consultancies.

These are *entirely* different purposes in nature from the third parties the DfE template privacy notice or DfE data usage page states it may share with which are organisations such as public bodies. [see page ¹⁴]

The department may decide to share pupil and children's information with third parties on a case-by-case basis where it is satisfied that to do so would be in accordance with the law and the [Data Protection Act \(1998\)](#), and where it considers that such disclosure would promote the education or well-being of children.

The [Education \(Supply of Information about the School Workforce\)\(No.2\)\(England\) Regulations 2007](#) allows us to share workforce data with the following third parties:

- [Local Government Association](#)
- [HM Chief Inspector of Education, Children's Services and Skills](#) (Ofsted)
- [Audit Commission](#)
- [Office of Manpower Economics](#)
- [Administrator of the Teachers' Pension Scheme](#)
- local authorities
- [Service Children's Education](#)
- school proprietors
- independent researchers where disclosure may be expected to be of public benefit

We only hold information provided to us by third parties such as schools, local authorities and awarding bodies. We do not generate any new information about individuals, and the data provided to us is only used to make decisions about educational policy on an aggregate statistical basis.

The data usage statement suggests "*the data provided to us is only used to make decisions about educational policy on an aggregate statistical basis.*"

There is no indication in this statement that sharing with third parties on a case-by-case basis is at individual level or identifiable.

¹⁴ <https://www.gov.uk/guidance/data-protection-how-we-collect-and-share-research-data>

There is no indication in this data usage statement or privacy template of commercial use or journalists use through and schools and parents and pupils cannot be reasonably expected to know this by any other means.

Commercial users and journalists working for newspapers for profit without evidence of public benefit do not meet any of the listed users.

The Department for Education has a legal duty which was reiterated by the CJEU in 2015¹⁵, to ensure fair processing before sharing personal data from the National Pupil Database with anyone else.

¹⁵ <http://curia.europa.eu/jcms/upload/docs/application/pdf/2015-10/cp150110en.pdf>