

## **Making Open Data Real: A Public Consultation Department for Education response**

We have consulted a wide range of stakeholders from across DfE, including policy officials, statisticians, legal, data protection and data security advisors. Overall we are very supportive of all the proposals, which align with our existing approach and aspiration to make even more DfE data more publicly accessible.

Our two main areas of concern are:

- i. **Balancing Transparency and Privacy** – ensuring that adequate safeguards exist to protect the privacy of individuals (especially children) for all published data and;
- ii. **Efficiency** – making sure that increased publication of data does not increase burdens on the resources of DfE or the front line, particularly schools.

Our responses to specific consultation questions are as follows:

### ***Where a decision is being taken about whether to make a dataset open, what tests should be applied?***

DfE already publishes a wide range of data about Education, including data about school performance and pupil attainment, school workforce and finance and Early Years and children.

For future decisions, we would want to include a test that:

- recognises that schools are small organisations, so new requirements impose cost and resource burdens and;
- recognises parents' concerns about use of data which could identify their children and/or lead to them being targeted for services and products

We also think that the potential utility and sustainability of making the data open on an ongoing basis should also be taken into account. The costs of making the data open should not outweigh the benefits it generates.

### ***If the costs to publish or release data are not judged to represent value for money, to what extent should the requestor be required to pay for public services data, and under what circumstances?***

DfE has already gone a long way in providing free access to its data: subject to legal and data protection considerations raw data can be requested free of charge by local authorities, government, academia etc. We have also initiated a programme of work to make even more Education data transparent and available.

We believe that data which is proactively made Open in this way, licensed for free reuse, will negate the need for any additional charging in the majority of cases. In fact, time spent responding to things like FOI requests will be reduced.

We would like to see:

- A cost/charging framework, similar to the existing three-days-work ceiling on FOI requests, to guard against vexatious requests.
- High standards for guidance and metadata, as in our experience more available data tends to generate more follow up questions.

We would also note that even if the requester is willing to pay additional costs, the extra work may prevent departments from delivering their core functions. There are also issues of fairness: should one requester gain greater access to data collected at public expense because they are able and willing to pay for extra analysis?

***How do we get the right balance in relation to the range of organisations (providers of public services) our policy proposals apply to? What threshold would be appropriate to determine the range of public services in scope and what key criteria should inform this?***

The viability or competitiveness of voluntary and community sector organisations should not be affected by potentially prohibitive costs met by trying to meet Open Data standards.

***What would be appropriate mechanisms to encourage or ensure publication of data by public service providers?***

Promoting the benefits of Open Data is key. DfE works closely with the front line (e.g. schools) and users (e.g. parents) to enhance efficiencies and communications, to achieve improved accountability, attainment, behaviour and attendance.

We would welcome mechanisms by which data providers are held accountable to public demand, where people know what data they are entitled to view, where to obtain it and its benefits.

***Meaningful Open Data: how should we ensure collection and publication of the most useful data, through an approach that enables public service providers to understand the value of the data they hold and helps the public at large know what data is collected?***

Like all providers of official statistics, DfE has robust internal procedures to ensure efficient collection of the most useful information. It is released with appropriate descriptions of purpose and quality, with any necessary caveats, to avoid accidental misuse or misrepresentation.

This practice should prevail with Open Data; data as collected may be fit for internal use but may not, if improperly described, be suitable for wider usage. Data could, and should, be reused, but proper governance and procedure should be in place to ensure data is viewed in its proper context.

For example, this is a key consideration in the work we are doing to provide improved access to a wider range of School Performance data than ever before. We are developing the Schools Performance tables to bring together for the first time school spending data, an enhanced set of school performance data, pupil cohort data and Ofsted judgements.

***Government sets the example: in what ways could we make the internal workings of government and the public sector as open as possible?***

DfE is highly supportive of public bodies and providers of public services proactively making data much more accessible. Education data is used by a wide range of users for a wide range of purposes and we have already taken large strides towards opening up access to our data sources, e.g. depositing research data in public archives for subsequent re-use, making statistical datasets available and proactively publishing all statistics and research findings.

***Are existing safeguards to protect personal data and privacy measures adequate to regulate the Open Data agenda?***

Protection of personal data is at the heart of the production of every DfE statistical release. Well established anonymisation processes are applied to the release of aggregated data (e.g. Neighbourhood Statistics, School Performance Tables, School Census) and we are working closely with relevant experts and stakeholders to fully assess the benefits and risks associated with disclosing pupil level information from our National Pupil Database.

We are concerned about the remote possibility of anonymisation being ‘cracked’ by motivated / well-resourced intruders, posing risk to the reputation of Open Data as well as possibly severe ramifications for the identified individual(s).

DfE would therefore seek assurances that the recommendations of the Cabinet Office’s review of privacy and transparency issues will be addressed.

***What might the resource implications of an enhanced right to data be for those bodies within its scope? How do we ensure that any additional burden is proportionate to this aim?***

It is a long standing DfE policy not to burden the front line or other data suppliers. Increases in transparency should not be at the expense of those suppliers, particularly schools. The costs of making the data open should not outweigh the benefits it generates. Proposals for an enhanced right to data should be accompanied by a full Impact Assessment.

***What is the best way to achieve compliance on high and common standards to allow usability and interoperability? Is there a role for government to establish consistent standards for collecting user experience across public services?***

DfE endorses the use of common standards to ensure usability and interoperability, and realise the benefits that Open Data can provide.

Careful consideration of how any data transparency standards are applied is required. We would be wary of any blanket “across government” regulations on the very specific points of practise where treatment on a system by system basis may be more appropriate. For example, different methods of anonymisation of personal data are applied to different DfE statistical releases to best maximise the protection offered depending on the precise nature of the data contained.

***How should public services make use of data inventories? What is the optimal way to develop and operate this?***

We would welcome the requirement for public service providers to have common, consistent and transparent data inventories to record what datasets are held, although we are mindful that this may not be the most accessible format for all users.

We would welcome guidance that is flexible enough to recognise that users of government data are wide ranging and have different needs.