

FINDASCHOOL RESPONSE (VIA E-MAIL)

I have recently set up a website (www.findaschool.info) designed to help inform parents' choice of primary school, and to help assess the chances of getting into a particular school. It therefore has a similar focus to schooloscope – mentioned in paragraph A1.15 (which unfortunately is now being closed down). findaschool.info is based entirely on the presentation of public data. I am therefore sharing my experiences in creating the site as a response to the consultation.

First and foremost, the move to encourage public sector data publication is extremely positive. From one perspective, it could be argued that the public sector ought to be providing services similar to that provided by findaschool. However, there are several advantages to services such as these being developed in the private sector. For example, relative to a public body a private company has:

- greater incentives to keep costs to a minimum;
- more scope to make mistakes and therefore greater flexibility to innovate and adapt;
- more opportunity to publish work-in-progress (i.e. beta versions).

The findaschool project started in June this year, with an early version of the website going live in September. It is still very much work-in-progress, and is improving all the time, but it has been providing a functional and useful service since its launch. It is clearly much more difficult for the public sector, and for large corporations, to operate in this way given the potential adverse reputational impact of mistakes.

The vast array of coding libraries that are now available mean that it is simpler and faster than ever to create online services. findaschool provides a simple example of what is now possible given access to the right data. However, the process would have been much faster, and therefore cheaper, if the source data were presented a more structured manner.

The main hurdle in developing the service has been collating the source data, and this remains the reason why the USP of the service – the ability to assess the likelihood of successful admission to a community primary school – is limited in scope to London and Manchester. This part of the service relies on data collected by local authorities regarding school admissions. Some local authorities publish the data, whereas others do not. Where they do publish, it is almost always in a non-machine readable format such as a table in a pdf document. Extracting this data and making it usable is time consuming and therefore very costly. Ultimately, this may prevent findaschool from being viable.

There are many possible solutions, but all relate in some way to specifying standards for the *structure* of public data – as discussed in section 8 of the consultation. A simple first step would be to require that public bodies publish source data in its raw format alongside formatted results. That is, whenever a document is published which includes tables, charts, or other data visualisations, the underlying data should be published separately in a format that can be more easily processed. Ofcom's Communications Market Reports provide a good example of this type of publication scheme. See for example, <http://stakeholders.ofcom.org.uk/market-data-research/market-data/communications-market-reports/cmrl1/telecoms-networks/>, which offers links to both pdf and csv versions of data tables.

UNCLASSIFIED

The precise nature of the data format is not so important. What matters is that is machine readable. As such, the benefit in moving from 1 to 2 stars (in the 5 star rating) is far greater than in moving up from 2 stars. The source data will almost always exist already in a machine readable format, and so it should not be too difficult or costly to implement such a publication scheme. Given the potential benefits in terms of the stimulus to developers, it seems reasonable to require that public bodies follow this advice when publishing data. If you have any further questions about this response, or findaschool.info, please let me know.

Regards

Ed Rushton

UNCLASSIFIED