

The SARs team at the University of Manchester are a small, ESRC funded, group committed to making microdata from the census as available as possible to as wide a range of users as possible. More information about the team and their work is available at <http://www.ccsr.ac.uk/sars>

We wish to comment on questions 1 and 2 of the consultation.

1. Do the definitions of the key terms go far enough or too far?

We did not find the definition of data to be very clear in the consultation document.

We are particularly concerned with individual-level data which are drawn from individual data subjects but subjected to anonymisation. These data are currently subject to licences which restrict their use. We were unclear from the consultation document whether such data are in scope.

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

The key tests should be:

1. is it legal to make these data open?

2. will making the data open improve or hamper the accessibility of the data?

With respect to the second of the tests, we seek to avoid any circumstance in which the utility of data currently available is reduced in order to make it suitable for release under an open government licence.

Consider a situation in which a micro dataset contains individual person records. These are routinely anonymised and may be released for analysis subject to the condition that the user does not attempt to de-anonymise those data. Anonymisation is done to respect the promise made to the data subject that their responses to the survey or census would be kept confidential and to avoid any harm to the respondent. This is important, particularly for data which are drawn from a census.

If these data can be confidently released under an OGL without reduction in detail we would very much welcome this as this would remove barriers to use.

However, should the detail of the data be reduced in order to make the data open, we would encourage data owners to consider alternatives to open data as a means of protecting the utility of the data.

The impact of data such as the Samples of Anonymised Records has been documented elsewhere (see <http://www.ccsr.ac.uk/sars/2011/documents/businesscase.pdf>) for example it has been shown that \$150 million of Inward investment into Northern Ireland was linked to analyses of these data. However, such use is contingent upon the data containing good multivariate detail. We seek to preserve that detail.