

'RSaBD' meeting, 2 Oct. 2015: Scoping paper SP9, Jeff Evans & Ludi Simpson

Big Data: information windfall, or merely hot air? *

At the time of writing, trends towards the availability of exceedingly large and continuously produced datasets have caught the imagination of a number of pundits and entrepreneurs. This sort of data, which requires ever larger storage facilities (previously unavailable or highly expensive), is produced by a range of sources: for example, monitoring devices operating continuously, such as CCTV security cameras or speed cameras; social media devices offering ample information on personal linkages, and attitudinal and behavioural data ('likes'); and consumption patterns, produced by supermarket 'loyalty cards'.

Such data pose a number of challenges to members of Radical Statistics and their allies:

Epistemological:

(a) a need to reassert the traditional concerns of data structure, quality, validity and bias, often played down by Big Data enthusiasts, in the belief that possessing a 'big' sample overrides such concerns, ignoring the fact that the sample may often be opportunistically selected (a striking example of the problem is given by the Literary Digest poll published before the 1932 US Presidential election (Marsh, 1979));

(b) continuing critique of the 'inductivist' belief that data alone can somehow generate dependable, general theories;

(c) reasserting the fact that 'big data' are often by-products of administrative systems and therefore both more affected by temporal changes of policy and procedure, and less revealing of social issues and tensions that current policy and procedures do not recognise.

Political-economic:

(a) challenging the monopolies of production and ownership of the facilities needed to produce 'big data';

(b) ensuring democratic control over the resources and methods of analysis for public policy.

Professional:

(a) developing skills in using the high quality computing facilities which are clearly going to become even more important in future;

(b) ensuring that the data analyst employees of big data owners gain and retain their key professional functions of quality control, independently of their employers' particular interests.

* Taken from: '**History and Role of the Radical Statistics Group**' [provisional title] (Jeff Evans & Ludi Simpson) ... chapter for Keiko Yasukawa & Stephen Black (Eds.) *Beyond Economic Interests: Critical perspectives in adult literacy and numeracy in a globalised world*; to be published by Sense Publishers, NL.

**** A copy is available for critical reading, on request. ****