

Is Big Data too big to fail?

Overview

The Americanisation of British politics continues apace. On Good Friday the Labour party excitedly announced, complete with a spelling error, that David Axelrod, President Obama's former Chief Campaign Strategist, had joined Ed Miliband's team for the 2015 general election. Axelrod will clearly bring with him an unenviable array of experience, talents and insights from his time helping elect Barack Obama; from the grassroots campaigning made infamous by the hitherto unheard of Democrat candidate in early 2007, to the vast array of data that helped secure his second term in 2012.

Axelrod's expertise lies in harnessing data, digital techniques and traditional doorstep canvassers to create a vast bank of potential voters. It was this belief in "cold hard data" which ensured his unwavering confidence, telling [CBS](#) on Election Day, that his candidate would be comfortably re-elected.

But what is big data? Why should we care? How is it used? And what, if any, impact will it have on British politics?

Flu trends

In its simplest form big data refers to the idea that "society can do things with a large body of data that weren't possible when working with smaller amounts". The term first pricked the public conscience in 2009 when a project, Google Flu Trends, claimed to have identified flu outbreaks from search queries alone. Released to much fan-fare, researchers have since discovered a problem with Google's Flu Trends system: it is no longer any good at predicting trends in flu cases. Indeed in the US it has wrongly forecasted the total number of cases for 100 of the past 108 weeks.

Industry and government

Companies such as Google, Amazon and Facebook pioneered the collection, processing and analysis of big data, whilst high-performance analytics is used across all sectors from banking and retail, to healthcare and insurance. Teams of analysts glean insights from data in just hours, minutes or seconds that once took days or weeks.



In engineering and manufacturing, for example, companies are finding new opportunities to predict maintenance problems, enhance manufacturing quality and manage costs using big data. In healthcare, there are new opportunities to predict and react more rapidly to critical clinical events, resulting in better care for patients and more effective cost management.

Even governments, including the UK, are jumping on the big data bandwagon. A [study](#) undertaken by SAS and the Centre for Economics and Business Research suggested that if the government capitalised on big data it could save £2bn in fraud detection, create 2,000 new jobs and generate £3.6bn in savings through better management of processes by, for example, integrating patient data to improve healthcare IT systems.

Formula One and Bergkamp

It is not just in the corporate world that big data is being utilised. In Formula One, as McLaren's cars speed round the track they send a stream of data back to the team that are processed and analysed in real time, enabling analysis of the data in seconds rather than hours. This real-time analysis of data helps the McLaren team make immediate proactive corrections, avoid costly, dangerous incidents, and ultimately win races.

Arsenal are known for their modern approach to the game, and their pioneering use of GPS to monitor their players in the early noughties. It was this early adventure which would explain why the Arsenal manager would so regularly substitute his flagship striker Dennis Bergkamp; the data simply showed he slowed down with twenty minutes to go and his performance duly suffered.



Hubris and protection

However, a consensus on big data, and its use, is far from settled. David Spiegelhalter, Winton Professor of the Public Understanding of Risk at Cambridge University, has described the enthusiasm around big data in technical terms as "complete bollocks. Absolute nonsense."

Critics central point is that big data's premise, in that a sampling bias can be overcome by simply sampling everything, is anything but comprehensive, and is even more prone to the statistical errors that haunt regular analytics.

Big data is also closely related to issues of data protection. This poses enormous challenges for industry, government, regulators and the public at large. Given the pace of the change there are concerns data protection legislation won't be able to keep up. It also ratchets up the importance of

