The Motley Coat of Data

Citation for published version:

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
The Motley Coat: The Motley Coat of Data

By Leslie Stevens

As a lawyer I am accustomed to, but sceptical of, hard and fast categories – we know all too well that bright line rules and seemingly fixed categories raise more questions than they provide answers. In data protection several such categories exist: anonymous data and personal data, sensitive personal data and ordinary personal data, data controller and data processor and so forth. Given the current reform of data protection law in Europe, and the anticipated introduction of the General Data Protection Regulation circa 2016/2017, questions should be (and have been) raised as to the effectiveness of these categories as regulatory tools. The lines between anonymous data and personal data have been blurred – the ease with which data may be re-identified and the power of data linkage has revealed the fallacy that is truly ‘anonymous’ data. Ordinary personal data, once considered innocuous, when combined with any countless array of other data, are likely to reveal intimate details about people’s lives – far more sensitive than data protection law would recognise.

My work for the ESRC funded Administrative Data Research Centre Scotland has led me to explore the issues raised by the use of administrative data in research (data which may involve health, but as a ‘category’, encompasses a far more vast range of areas including education, tax, benefits, housing, transport and so forth). I have questioned what meaningful differences – if any – exist between the governance of administrative data and health data. While the data landscape is admittedly diverse and a motley coat of sorts, what lessons can be learned and principles adopted from the governance of one type for another? I have found that it is what these two seemingly dissimilar ‘categories’ have in common that is far more useful in discerning what good governance looks like in the administrative data context – the importance of the public interest to support certain uses of data.

Again, as a lawyer, I am all too familiar with (and quickly frustrated by) the deployment of impossibly vague terms in the law, including the public interest. However, in response to the term’s constant use in the research context (used to justify new and ever expanding uses of personal data – albeit ‘anonymised’) and the lack of solid guidance on its appropriate use and meaning in data protection, I became fascinated by the concept and resolved myself to dig deeper into the murky world of the public interest. Through exploring this varied and venerable heritage of the public interest I have become more convinced of its importance as a tool and as a facilitator for important debates over what can justifiably, lawfully and ethically be considered acceptable uses of data in a particular context. The public interest has become my theoretical and practical tool for understanding the fundamental tensions at work in the motley coat of data.

Analysed through the prism of the public interest it becomes clear that the ever-shifting balance between privacy and promoting the lawful and ethical use of data for research is not only delicate but is never as black and white as it may first seem. Balancing risks to privacy versus any potential loss to the public interest in terms of socially/economically/culturally
valuable research not undertaken involves multiple public interests and numerous 'publics' - there are no simple 'yes' or 'no' answers. The concept of the public interest forces me to accept the inherently varied nature of the data protection landscape – there is no 'single' public interest but rather an ever-shifting balance that must be considered in light of the particular context in question.