

How do public values get into public sector machine learning systems, if at all?*

Michael Veale
University College London
Gower St, London WC1E 6BT
m.veale@ucl.ac.uk | @mikarv

November 15, 2016

Many areas of concern around machine learning that have been raised in academia, the media or by civil society concern public sector applications. Despite this, little empirical work has been done on understanding how the public sector currently obtains, pilots, deploys and manages these technologies. While considerable strands of work are emerging on social and technical approaches to issues like fairness, transparency and accountability, it remains strongly divorced from on-the-ground realities. This divorce seems certain to hinder applicability and uptake, mimicking the non-adoption of sophisticated privacy-enhancing technologies (PETs) designed with particular sectors and issues in mind.

This study attempts to improve the currently largely absent evidence base around the challenges of using machine learning technologies responsibly in practice. Interviews with over thirty individuals involved with machine learning in and with public sector organisations in five countries were carried out in order to understand actors' experiences with machine learning systems in their own words. The areas explored include taxation, child protection, police and other emergency services, prison management, border control and local government. Individuals ranged from civil servants responsible for particular projects, to the contractors and developers tasked with delivering them. Concepts raised in these interviews were grouped and analysed using a typology of public sector values from the public administration literature, providing a structured overview as well as placing aired machine learning concerns in a broader context.

Preliminary results¹ indicate that i) the common assumption that future populations are not functions of past decisions is often violated in the public sector; ii) divergent and often new internal governance mechanisms attempt to deal with specific machine learning issues iii) while academics focus on the issues around values in models, public sector organisations tend to problematise the values emerging from decision-support delivery systems instead. The full paper also contains a range of topics and issues raised in these interviews in order to open up the debate with diverse examples, rather than close it down with generalisations.

*This work has been supported by the UK's Engineering and Physical Sciences Research Council (EPSRC) through a Doctoral Training Partnership grant (EP/M507970/1).

¹The above wave of interviews has only recently concluded.