**Biography**

Jake Goldenfein is a law and technology scholar interested in the emerging structures of governance in computational society. He completed a PhD at Melbourne Law School (University of Melbourne) and has been a lecturer at Swinburne Law School since 2016. His research addresses: surveillance, automation and identity; law in cyber-physical systems; platform law and governance; the role of standards in law-making; distributed ledgers and intellectual property; media history and theory; and legal theory in the context of mega-computation. His upcoming book ‘Monitoring Laws: Surveillance, Automation and Information Law’ will be published by Cambridge University Press in 2019.

**Abstract**

Over the past decade, researchers have been investigating new technologies for categorising people based on physical attributes alone. Unlike profiling with behavioural data created by interacting with informational environments, these technologies record and measure data from the ‘real world’ and use it to make a decision about the ‘world state’ – in this case a judgement about a person. Automated Personality Analysis and Automated Personality Recognition, for instance, are growing sub-disciplines of computer vision and computer listening. This family of techniques has been used to generate personality profiles, assessments of sexuality, political orientation and criminal propensity using facial morphologies and speech expressions alone. These profiling systems do not target the content of images or speech, but measure and analyse para-visual and para-sonic information to train classifiers for revealing non-visual information like personal typologies and behavioural predictions.

While the knowledge claims of these profiling techniques are often tentative, they increasingly deploy a variant of ‘big data epistemology’ suggesting there is more information in a human face or in spoken sound than is accessible or comprehensible to humans. This paper explores the bases of those claims and the systems of measurement that are deployed in computer vision and listening. It asks if there is something new in this class of data science knowledge claim, and attempts to understand what it means to combine computational empiricism, statistical analyses, and probabilistic representations to produce knowledge about people. Finally, the paper explores possible mechanisms for contesting the emergence of computational empiricism as the dominant knowledge platform for understanding the world and people within it.