# Applying 'Big Data' Privacy Concepts to Predictive Phenotyping

Nathan Scudder[1 2], Dennis McNevin[2], Sally Kelty[3], Simon Walsh[1], James Robertson[2]

[1] Australian Federal Police, GPO Box 401, Canberra ACT 2601, Australia
[2] National Centre for Forensic Studies, Faculty of ESTeM, University of Canberra, ACT 2617, Australia
[3] Faculty of Health, University of Canberra, ACT 2617, Australia

Big Data approaches have been applied extensively in business, but the technology is advancing in medical science and genetics.[i] Big Data concepts are already applied to clinical data and translational research.[ii] These approaches can also apply to forensic use of predictive phenotyping, and this use can also inform Big Data.

Big Data, genetic or otherwise, often lacks meaning without context and can also include extraneous data. The potential for data to be contextualized challenges conventional approaches to de-identifying personal information.[iii]

Technology often drives how privacy is defined.[iv] Phenotyping, by its nature, seeks to add identifying context to genetic data. Successful use of forensic phenotyping is therefore undertaking the very re-identification that is challenging to Big Data privacy more generally. Examining phenotyping through a Big Data privacy and governance lens can provide useful insights to forensics and overarching Big Data.

---

[i] Moses, Lyr a Bennett and Janet Chan, 'Us ng b g data for ega and aw enforcement dec s ons: Test ng the new too s' (2014) 37 UNSWLJ 643 http://www.unsw aw ourna .unsw.edu.au/s tes/defau t/f es/t1 bennett moses and chan.pdf, 650; Chute, Chr stopher G et a , 'Some exper ences and opportun t es for b g data n trans at ona research' (2013) 15(10) Genet cs n Med c ne 802 http://www.nature.com/g m/ ourna /v15/n10/fu /g m2013121a.htm .

[ii] Chute, above n .

[iii] Mayer-Schönberger, V ktor and Yann Padova, 'Reg me Change? Enab ng B g Data through Europe's new Data Protect on Regu at on' (2016) 17 Co um. Sc . & Tech. L. Rev. 315 http://st r.org/down oad/vo umes/vo ume17/SchonbergerPadova.pdf, 328; Anderson, Nate, '"Anonym zed" data rea y sn't - and here's why not (B g Data)' (2009) http://arstechn ca.com/tech-po cy/2009/09/your-secrets- ve-on ne- n-databases-of-ru n/; Kempe, Shannon, 'Best Pract ces for B g Data Governance' (2013) http://www.datavers ty.net/best-pract ces-for-b g-data-goverance/.

[iv] Wh te House, 'B g Data and Pr vacy: a techno og ca perspect ve' (2014) *Washington DC: Executive Office of the President President's Council of Advisors on Science and Technology* https://www.wh tehouse.gov/s tes/defau t/f es/m cros tes/ostp/PCAST/pcast b g data and pr vacy - may 2014.pdf, 3-4; P ke, E zabeth R, 'Secur ng Sequences: Ensur ng Adequate Protect ons for Genet c Samp es n the Age of B g Data' (2015) 37 *Cardozo Law Review* http://papers.ssrn.com/so 3/papers.cfm?abstract d=2658306, 33.