Edwards, A. (2015) 'Big Data, Predictive Machines and Security: Enthusiasts, Critics and Sceptics', *Discover Society*, on-line publication at: http://discoversociety.org/2015/07/28/big-data-predictive-machines-and-security-enthusiasts-critics-and-sceptics/

Adam Edwards
Cardiff School of Social Sciences
Email: EdwardsA2@cardiff.ac.uk

Big data is in part responsible for a rejuvenated interest in the use of artificial intelligence for security applications such as predictive policing. Enthusiasts argue that using this data in concert with advances in computational machine learning will enhance the capacity to anticipate and pre-empt problems of crime and security[1]. In response, critics either challenge the political feasibility of predictive machines, noting their unjustifiable intrusion into private communications or their self-fulfilling prophesies, or they doubt the technical feasibility of training machines to learn the improvisation that is central to social relations such as crime and insecurity. Alternatively, sceptics are concerned with the powers and liabilities of hybrid human-machine learning in predicting security 'scripts'.

Enthusiasts

The emergence of 'social computing' given the arrival of 'read/write technologies' (such as blogs, micro-blogs, social networking etc.) on the 'interactive World Wide Web' has provoked further interest in the prospect of integrating such user-generated data with digitally collected and archived administrative and commercial datasets to provide a 'bird's eye view' of social relations. Computational scientists refer to this amalgamation of big data in terms of the '10,000 foot view' of 'the social graph' and argue that read/write technologies are just the beginning of an 'age of social machines'[2]. It is envisaged that these machines will rapidly evolve from a situation in which various read/write applications operate in isolation from one another (an exchange on Facebook, a discussion on Twitter, comments on a broadcast media website, opinions registered through on-line surveys, the retrieval and annotation of digitally archived police, health, education and census data and so forth) to one in which they interact with each other. In this way it is believed that social machines will enable an exponential increase in the kinds of artificial intelligence and collaborative work needed to grasp and solve the complexity of social problems that confront us and which are irremediable through individual thought and effort, from climate change through major public health challenges to mobilising local community responses to crime and violence.

[1] Vlahos, J. (2012) 'The Department of Pre-Crime', *Scientific American*, 306/1: 1-9.
[2] Hendler, J. and Berners-Lee, T. (2010) 'From the Semantic Web to social machines: A research challenge for AI on the World Wide Web', *Artificial Intelligence*, 174: 156 – 161.

There is fertile speculation about the kinds of security scenarios that could unfold once such faith is placed in artificial intelligence. Enthusiasts for predictive policing (PREDPOL)[3] in the United States and for 'prospective crime mapping' (PROMAP)[4] in the United Kingdom have developed algorithms, premised on a 'contagion thesis', which seek to detect when and where crimes will occur by factoring in different kinds of assumptions about how crime spreads from an initial offence in particular environments given the routine activities and rational calculations of offenders, victims and control agents[5]. These predictions are then tested against the crime patterns actually registered through conventional methods of police recording and self-report studies of offending and victimisation and then the algorithms are subsequently revised as a means of better anticipating crimes and targeting pre-emptive interventions. The enthusiasm for building predictive machines is now being further extended to design algorithms or 'machine classifiers' to better 'sense' and anticipate patterns of threatening or 'hateful' on-line communications through social media and forecast their putative relationship to off-line events such as terror attacks[6].

Critics

There are predictable criticisms of engineering predictive machines for security applications. In this context the most notable are those political concerns raised by the whistle-blower Edward Snowden about the massive and routine invasion of privacy through the US National Security Agency's PRISM surveillance programme. PRISM collects communications through the internet, without reasonable suspicion, and then mines them for intelligence on, and forecasting about, various security threats including terror plots and illicit drugs markets[7]. The Snowden revelations suggest how an understanding of security premised on Big Data necessarily contravenes the right to private communications because the 10,000 foot view of the 'security graph' cannot be accomplished without generalised data collection from whole populations.

A related concern is that predictive machines generate self-fulfilling prophesies. They become active ingredients in the targeting of suspects such that problems of security become artefacts of the way the algorithms, machine classifiers and underlying assumptions of predictive machines, speculate about security to include certain concerns (e.g. speech about 'radicalised' Muslim youth) whilst obviating others (e.g. speech about the culpability of Western foreign policy in the Middle East). In this regard, targeting the usual suspects ceases to be just a consequence of episodic prejudicial police actions and becomes automatically reproduced by a social machine. The alienation of entire social groups as a consequence of this kind of group profiling and targeting, along with the creation of a policing environment conducive to miscarriages of justice is an established theme in critical criminology, particularly in the UK with reference to the war in Ireland and the long history of antagonism between the police and street populations, particularly of young males from minority

[3]Perry W.L., McInnis B., Price C.C., Smith S.C., Hollywood J.S., 2013, *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*, Rand Corporation Report.

[4] Johnson S., Bowers K., Birks D., Pease K., (2009) 'Predictive Mapping of Crime by ProMap: Accuracy, Units of Analysis, and the Environmental Backcloth', in D. Weisburd (ed.), *Putting Crime in its Place*, New York, Springe.

[5] Benbouzid, B. (2015) 'From situational crime prevention to predictive policing', *Champ pénal/Penal field*, Vol. XII: URL : http://champpenal.revues.org/9066; DOI : 10.4000/champpenal.9066.

[6] Burnap. P. and Williams, M.L. (2015) 'Cyber Hate Speech on Twitter: An Application of Machine Classification and Statistical Modeling for Policy and Decision Making', *Policy and Internet*, 7/2: 223-242.

[7] Lyon, D. (2014) 'Surveillance, Snowden, and Big Data: Capacities, consequences, critique', *Big Data and Society*, July-December 2014: 1-13.

social groups[8]. However, the consequences of automating this policy failure are only just beginning to be appreciated[9].

Finally enthusiasm for predictive machines can be criticised on grounds of technical feasibility. It is argued there are only a few instances in which machines can be programmed to effectively mimic human actions, such as swinging golf clubs or dialling telephone numbers, whilst there are many of these actions that machines cannot accomplish and, crucially, will never be able to accomplish, because they require the fundamentally human capacity for improvisation, such as in writing love letters or subverting factory work regimes[10]. From this perspective, the potential success or colossal failure of predictive policing hangs on the question of how improvised security problems are. It has been argued, for example, that improvisation is the central dynamic of much crime, particularly sophisticated organised crimes, in which perpetrators and preventers are in an ongoing correspondence, in this case an 'arms race' rather than an amorous exchange, to outflank and outwit each other[11]. In this scenario predictive machines will fail because they cannot adapt quickly enough to improvised real world social relations.

Sceptics

Sceptics of both the enthusiastic embrace and rejection of predictive machines argue that social relations may not be akin to swinging golf clubs but they may be sufficiently 'scripted' to be predicted, in part, by automated learning. Security problems may, to continue the analogy, be more like performances of a play in which the actors improvise around the script but still rehearse their lines and do not completely rewrite the story from one performance to another. How these scripts and their narrative structures can be understood is the subject of current methodological argument and innovation in 'digital social research', in particular the prospects for hybrid human-machine learning in which algorithms driving the automated collation and analysis of big data are collaboratively designed and frequently refreshed [12].

[8] Pantazis, C. and Pemberton, S. (2009) 'From the "Old" to the "New" Suspect Community: Examining the Impacts of Recent UK Counter-Terrorism Legislation', *British Journal of Criminology*, 49: 646-666; Hallsworth, S. and Lea, J. (2011) 'Reconstructing Leviathan: Emerging contours of the security state', *Theoretical Criminology*, 15/2: 141 – 157.

[9] Chan, J. and Bennett Moses, L. (2015) 'Is Big Data challenging criminology?', *Theoretical Criminology*, May 19, 2015 1362480615586614.

[10] H. M. Collins and M. Kusch (1998) *The Shape of Actions: What Humans and Machines Can Do*, Cambridge MA, The MIT Press, p. 31.

[11] Ekblom, P. (2003) 'Organised Crime and the Conjunction of Criminal Opportunity Framework', in A. Edwards and P. Gill (Eds.) *Transnational Organised Crime: Perspectives on global security*, London, Routledge; Dorn, N. (2003) 'Protieform criminalities' in A. Edwards and P. Gill (Eds.) *Transnational Organised Crime: Perspectives on global security*, London, Routledge.

[12] Edwards, A., Housley, W., Williams, M.L. et al (2013) 'Digital Social Research, Social Media and the Sociological Imagination: Surrogacy, Augmentation and Re-orientation', *International Journal of Social Research Methodology*, 16/3: 245-60; Housley, W, R Procter, Edwards, A., et al. (2014) 'Big and broad social data and the sociological imagination: a collaborative response', *Big Data & Society* April–June 2014: 1–15.

An example of a security script is Hope's thesis on the 'immunisation' of non-victims of volume crimes, such as domestic burglary in England and Wales[13]. In this script, non-victims immunise themselves through their access to private 'club goods' such as access to commercial household security, the market value of their homes as a proxy for the segregation of the residential population into wealthier neighbourhoods less susceptible to crime and through their participation in 'gated communities' with enhanced security surveillance and patrols. It is argued this immunisation explains the grossly unequal distribution of victimisation and non-victimisation for this problem in which four fifths of the residential population experience only one fifth of the burglary whilst 20 per cent of the population, with negligible access to these security club goods, are chronically victimised, experiencing an estimated 80 per cent of the problem. There may be improvisation around this script, in that property crime is often further concentrated within particular households even within 'high crime' neighbourhoods, but this is unlikely to result in any radical redistribution of household burglary by neighbourhood and therefore any refutation of the immunisation script. This is an institutional question of the balance between public and private policing rather than a question of household dynamics. If this distribution of victims and non-victims is relatively stable, it ought, therefore, to be anticipated by a predictive machine driven by algorithms premised on Hope's thesis. Whether more improvised security problems, such as urban riots, could ever be anticipated by predictive machines is a moot point but between the extremes of stable domestic burglary patterns and episodic civil unrest, there are other security problems whose scripted and predictable qualities are open to further research, debate and argument. An example of this is the use of artificial intelligence to indicate tension in social media communications[14].

Sceptics consequently offer an altogether messier and less certain reflection on the limits to hybrid human-machine learning but one that is irreducibly driven by humans in constituting problems of security not simply registering objective truths. There are also grounds for scepticism about the integration of variegated data sets composed of material collected over hugely varying temporal and spatial horizons. There may be opportunities for recomposing this data in ways that enable it to be meaningfully linked but even where relatively robust administrative data sets are concerned this entails a substantial input from human intelligence. Whether and how the lower fidelity data generated by users of social media can be meaningfully linked to other administrative and commercial data as well as the primary data sets produced by social scientists remains a very challenging, possibly insurmountable, methodological problem. Even if technically feasible, and this is a very big 'if', there are genuine ethical and political concerns about engineering predictive machines capable of collating person-specific data from multiple sources in order to circumvent controls on the anonymity of such data, and thus the profiling and monitoring of 'risky' individuals and groups. This is particularly so where predictive machines could be deployed within any governing regime not just in liberal democracies with lively, open, debates about 'snooper's charters' and Orwellian objections to Big Brother. At best we are in a situation that requires

[13] Hope T. (2006) 'Mass consumption, mass predation - private versus public action? The case of domestic burglary in England and Wales', in Lévy R., Mucchielli L., Zaubermann R. (Eds.), *Crime et insécurité : un demi-siècle de bouleversements. Mélanges pour et avec Philippe Robert*, Paris, L'Harmattan, 46-61.

[14] Williams, M.L., Edwards, A., Housley, W. et al (2013) 'Policing Cyber-Neighbourhoods: Tension Monitoring and Social Media Networks', *Policing and Society*, 23/4: 461 – 481; Burnap, P., Rana, O., Avis, N. et al (2015) 'Detecting Tension in On-line Communities with Computational Twitter Analysis', *Technological Forecasting and Social Change*, 95: 96-108.

deliberation about the levels of confidence inspired by the technical feasibility of predictive machines and then about the appropriate regulatory frameworks for governing the access to and uses of such machines.