Norman Stanley Fletcher and the case of the proprietary algorithmic risk assessment

Marion Oswald (Winchester University) and Jamie Grace (Sheffield Hallam University)

Norman Stanley Fletcher, the flawed hero of the BBC's prison sitcom 'Porridge' could have been in no doubt of the factors which had resulted in the decision to send him to prison for five years. At the start of each programme, the judge's voice boomed:

> 'You have pleaded guilty to the charges brought by this court…You are an habitual criminal, who accepts arrest as an occupational hazard, and presumably accepts imprisonment in the same casual manner. We therefore feel constrained to commit you to the maximum term allowed for these offences.'

That was the 1970s. In 2016, another factor may be added to the mix, in the US at least: the ***proprietary*** algorithmically generated risk assessment.

Loomis decision – Supreme Court of Wisconsin

The Supreme Court of Wisconsin has recently ruled on the challenge by convicted criminal Eric Loomis to the use of an algorithmic risk assessment tool called 'COMPAS' during his sentencing ('Correctional Offender Management Profiling for Alternative Sanctions' owned by a company called Northpointe). COMPAS and similar tools were initially used only by probation and parole departments but have now expanded to sentencing. As Brustein, writing in Bloomberg, commented 'using algorithms to decide how long someone goes to jail is proving more controversial than using them to decide when to let people out early.'

The COMPAS risk assessment is based upon information gathered from the defendant's criminal fil and an interview with the defendant. It generates risk scores representing pre-trial recidivism risk, general recidivism risk, and violent recidivism risk. As the judgment explained:

> 'risk scores are intended to predict the general likelihood that those with a similar history of offending are either less likely or more likely to commit another crime following release from custody. However, the COMPAS risk assessment does not predict the specific likelihood that an individual offender will reoffend. Instead, it provides a prediction based on a comparison of information about the individual to a similar data group.'

Loomis asserted that the court's consideration of a COMPAS risk assessment at sentencing violated his right to due process. In particular he argued that because the tool was based on group data, it violated his right to an individualised sentence (the tool was designed to predict group behaviour); and the proprietary nature of the tool prevented defendants from challenging the assessment's scientific validity. The court ultimately dismissed Loomis's claims, thus allowing a COMPAS risk assessment to be used at sentencing. The court overcame the concerns about group data, commenting (perhaps rather optimistically):

> 'Just as corrections staff should disregard risk scores that are inconsistent with other factors, we expect that circuit courts will exercise discretion when assessing a COMPAS risk score with respect to each individual defendant.'

It dismissed Loomis's argument that unless he could review how the factors are weighed and how risk scores are determined by the tool, the accuracy of the assessment cannot be verified. The court stated:

> 'Although Loomis cannot review and challenge how the COMPAS algorithm calculates risk, he can at least review and challenge the resulting risk scores set forth in the report'

further declaring that Loomis had an opportunity to challenge his risk scores by arguing that other factors or information demonstrated their inaccuracy.

The court did however acknowledge that a number of studies had questioned whether the tool was actually assessing recidivism risk accurately. In addition, concerns had been expressed that the tool may disproportionally classify ethnic minority offenders as higher risk (according to an analysis done by ProPublica subsequently denied by Northpointe). Therefore, risk scores must not, the court said, be used to determine whether an offender is incarcerated or the length of the sentence, and should not be determinative of whether an offender can be supervised safely in the community. The tool should additionally be subject to a number of cautions that a court should take into account as follows:

- the proprietary nature of COMPAS prevents disclosure of information relating to how factors are weighed or risk scores are determined;
- the scores are based on group data, and so are not able to identify a particular high-risk individual;
- concerns have been raised about disproportionate classification of ethnic minority offenders as high risk;
- the scores are based on a national sample; there had been no cross-validation for a regional/State sample;
- COMPAS was not developed for use at sentencing.

UK context

The UK government has for some years now invested in the use of database-driven means of predictive tools for offender management purposes in the prison and probation contexts, with the Offender Assessment System reoffending and violence predictor scores used not only in parole and probation contexts, but also in pre-sentencing reports. The National Offender Management Service's 2015 evaluation sets out a commendably transparent review of the validity and reliability of the current system, noting that the system required continuous evaluation and recalibration based on publicly observable scoring rules and definitions. The evaluation also assessed actual/predicted reoffending differences by gender and ethnicity, concluding that lower validity for minority ethnic groups was the greatest concern (so requiring remodelling of predictors and responsible use by offender managers).

In the wider criminal justice context, our recent FOI-based study of all national police forces indicated that only a relatively small percentage of forces (14% of those who responded, representing 6 forces) use algorithmic tools to analyse ***intelligence*** data (as opposed to committed crime data). One force reported that that their systems were created by private companies and then modified for use within the force. Decision-making and risk assessment relating to individuals were included in the reasons for using such algorithmic tools. However, there was no clear indication in the FOI responses as to the specific activities, crimes, schemes or laws that were the focus of these tools.

Are, for example, algorithmic tools used in the application of the Domestic Violence Disclosure Scheme (known as 'Clare's Law')? This is a scheme under which the police have to decide whether to supply public protection risk information, based on an actuarial judgement. For schemes where difficult risk-based judgements are required, it is easy to see why a *reliable* algorithmic decision-making tool would be potentially helpful, provided not used in a determinative way. However, if the implementation of such tools resulted in less human judgement and more machine-driven decisions - especially if driven by proprietary, less than transparent, algorithms - natural justice and procedural fairness claims may well arise.

The key House of Lords decision in *Doody* [1993] UKHL 8 (a challenge to the Home Secretary's decision in setting the tariff for life sentence prisoners and thus departing on occasions from the view of the judiciary) makes a number of pertinent points relating to the duty to give reasons in a sentencing context. Lord Mustill dismissed the Home Secretary's argument that the prisoner can deduce without the need for any more information both the factual basis of the Secretary of State's decision, and the intellectual reasons why the penal element was fixed at a particular term of years:

> 'The prisoner does indeed know what primary materials were before the court, but he does not know what the judge and the Home Secretary made of them, nor does he know what other materials, not brought out at the trial, may have formed an element in the decision.'

Lord Mustill went on to comment [emphasis added]:

> 'Where a defendant is convicted of, say, several armed robberies he knows that he faces a stiff sentence: he can be advised by reference to a public tariff of the range of sentences he must expect; he hears counsel address the judge on the relationship between his offences and the tariff; he will often hear the judge give an indication during exchanges with counsel of how his mind is working; and when sentence is pronounced he will always be told the reasons for it…Contrast this with the position of the prisoner sentenced for murder. **He never sees the Home Secretary; he has no dialogue with him: he cannot fathom how his mind is working**.'

It would seem unlikely therefore that the UK Supreme Court would allow the secret workings of a proprietary risk-assessment algorithm to form part of a court's sentencing deliberation as the Wisconsin Supreme Court has done. It is to be hoped that, as the *Doody* decision held that a Home Secretary must show 'how his mind is working', the same will be true of an algorithm. In relation to public protection disclosure schemes involving risk-based decisions, again it seems unlikely that commercial confidentiality would be permitted to be a barrier to appropriate scrutiny. This suggests that those drafting public sector procurement contracts with third party software suppliers should require a certain level of transparency from the commercial party, including disclosure of the algorithmic workings in a way that would facilitate investigation by a third party in an adversarial context if necessary. Accountability and transparency requirements might even lead towards the conclusion that open source algorithms should be used by default.

Finally, we believe that now is the time to have an open and informed debate about the use of algorithmic tools across all aspects of public life so that society can decide where the red lines should be. Learning lessons from the recent Investigatory Powers Bill debate, the legal, ethical and policy issues surrounding the use of such tools need to be tackled transparently and without delay. As Harkness comments:

> 'We have to make decisions, based on too little information, with unforeseeable consequences. No wonder people are tempted to let a machine make the predictions, take the decisions, and bear the blame.'

Should that temptation always be resisted or are there circumstances in which a machine-based process would genuinely produce fairer and more consistent decisions? How can we ensure that supposedly advisory algorithmic assessments do not in practice have undue influence and so become the default decision-maker? An ability to critically assess the way in which an algorithm's 'mind is working' will be crucial to these and no doubt the many more questions that will arise in the future.

Save the date! The 4th Winchester Conference on Trust, Risk, Information & the Law (TRILCon17) will be held on Wednesday 3 May 2017, focussing on de-personalised decision-making, machine-

power, A.I. and drones. Keynote from Professor Katie Atkinson, Head of Computer Science, University of Liverpool, an expert in A.I. and its application to judicial/legal reasoning.