In: Emre Bayamlıoğlu, Irina Baraluic, Liisa Janssens and Mireille Hildebrandt (eds), BEING PROFILED:COGITAS ERGO SUM. 10 Years of Profiling the European Citizen, 2018: Amsterdam University Press, 58-61. DOI 10.5117/9789463722124/CH9.

The presumption of innocence's Janus head in data driven government

*Lucia M. Sommerer**

Abstract

This provocation argues for a broader interpretation of the presumption of innocence: a presumption of innocence that applies to government actions not just during, but also prior to, the criminal trial. Notably, it argues for the application of a special standard of certainty not just for criminal convictions (reasonable suspicion), but also for the application of a similarly uniform standard to data driven, algorithmic risk-assessments which trigger actions such as pre-trial detention. At a minimum, this provocation hopes to prompt an overdue discussion of the number of false positives our society is willing to accept in our pursuit of security. By examining the example of the pre-trial risk-assessment tool COMPAS, this provocation points out how statistical consumer profiling adapted from the advertising industry is migrating to the criminal justice sector, and how the nuances of statistical likelihoods are ironed out within the criminal justice context and molded into "legal truth". Further, it stresses the risks of algorithmic decision making that is gradually colonizing many aspects of the criminal justice system. In particular, it draws attention to a possibly problematic focus shift from concretely defined criminal actions to criminally relevant behaviour and attitudes in substantive criminal law. It concludes by inviting the reader to imagine the lenses through which future generations may one day – amused or reproachful – look back on our analyses of data driven government.

Keywords: presumption of innocence, new penology, 'Lifestyle-Guilt', risk assessment, recidivism prediction

Introduction

> '*Contemporary practices of risk operate in a way that precludes the possibility of a non-dangerous individual.*'
>
> Werth (2018, 1)

One reason for the existence of the presumption of innocence (PoI) is to prevent premature, wrongful convictions. Criminal convictions are one of the most serious ways in which a state can diminish its citizens' freedom. Imagining having to endure a conviction as an innocent person triggers a collective fear. Those wrongfully convicted despite procedural safeguards often describe their experience with such phrases as, 'I felt buried alive!'. If no standards existed to ensure that in principle no innocent person is subject to a criminal conviction, we might live in constant fear: a wrongful conviction could happen to any one of us at any time. The PoI sought to remedy this fear by imposing procedural standards on criminal trials. The question that shall be posed here, however, is: in a data driven government is it sufficient to remedy the threat of wrongful convictions by applying the traditional (narrow) PoI? Or is a broader interpretation of the PoI warranted because the contexts in which an innocent person may be 'buried alive' are no longer limited to criminal trials, but have expanded substantially into the time before a criminal trial. General arguments in favour of a broader PoI have been made in the past (Duff 2013; Ferguson 2016), but in this provocation I will focus on the possibly intrinsic need for a broader PoI in data driven governments.

Two readings of the PoI

Traditionally the gaze of the PoI was turned towards the past, i.e. to prior criminal actions. It has applied in cases where someone was accused of having committed a crime, and in the subsequent criminal proceedings. This narrow PoI is thus trial-related, which means it is related to limiting state actions taken only after an alleged criminal act.

A broader reading of the PoI should be adopted. This means a reading not limited to the criminal trial, but instead also related to risk assessments, i.e. to the suspicion that someone will commit a crime in the future. Such a broad reading is needed under data driven governments in which algorithmic profiling of individuals is increasingly used to determine the risk of future criminal activity, and where the criminal justice system attaches materially negative consequences to an individual's high-risk score. Based on pattern matching, it is determined whether someone belongs to a risk group that warrants the attention of the criminal justice system. Statistical profiling, a technique borrowed from behavioural advertising, has increasingly infiltrated the criminal justice sector, along with its existing shortcomings (opacity, discriminatory effects, privacy infringements, false positives) while also creating new ones, notably the risk of prejudgment.

The PoI should thus be a guiding principle not just for the repressive branch of the criminal justice system but must also be applied to preventive pre-trial decisions. Such an extension would provide the PoI with a protective, Janus-faced gaze into the past and into the future simultaneously. This gaze comes in the form of a special standard of certainty required for both a criminal trial conviction and a high-risk determination triggering pre-trial detention.

A classification of high-risk generally does not claim to predict behaviour with near certitude. Rather, it claims that an individual shares characteristics with a group of people in which higher levels of criminal activity are present compared to the rest of the population. As an illustration, consider the pre-trial risk-assessment tool COMPAS, which is used in bail decisions in the U.S., and which equates an 8% likelihood of being arrested for violent crime in the future with high-risk status (Mayson 2018). A high-risk individual under COMPAS thus shares characteristics with a group of people of which 8% have been rearrested for a violent crime in the past. The nuances of this statistical judgment, including the lack of certitude are, however, ironed out in the application of the statistics. The likelihoods turn into “legal truth” for defendants when a judge at a bail hearing is presented with a high-risk classification (which generally neglects to mention the underlying statistics), and when defendants as a direct or partial consequence are then denied bail. The statistical information that out of a group of say 100 people with the same high-risk characteristics as a defendant only 8 have committed violent crimes turns into the “legal truth” that the defendant is dangerous and must be monitored further through denial of bail. The possibility that the individual belongs to the 92 individuals that have not committed a crime is not considered: the possibility of a non-dangerous individual in this group is, to use Werth’s wording, precluded.

In this context one can speak of a dormant penal power embodied in the various data collected and in the profiles created about individuals. This penal power comes to life once a profile is fed into criminal justice algorithms that generate risk scores, which in turn are used as basis for a suspicion or as a justification for further criminal justice measures.

I argue that a mistake (false positive) in this risk context (A is found to be high-risk even though he is not), and in the context of a trial (B is found guilty even though he is innocent), are mistakes of the same kind. In both situations individuals receive a treatment they do not deserve. In the trial scenario, the PoI requires a special standard of certainty (beyond reasonable doubt) to convict someone in order to prevent wrongful convictions. A broad PoI would also require such special, uniform standard of certainty for ranking an individual high-risk and attaching manifestly negative criminal justice consequences (e.g. pre-trial detention) to this risk score. Similarly, a special standard of certainty would also be required for individual-level risk-assessments deployed by law enforcement. The standards in these cases may not be beyond reasonable doubt, but they would most certainly require more than 8% likelihood, and they would at least elicit an overdue public debate over what percentage of false positives society is willing to tolerate in its pursuit of security.

These described algorithmic developments are at the ridge of a deep conceptional shift – the so-called new penology – that appeared in criminal justice in the late 20th century. This shift was marked by the emergence of a then-new discourse on probability and risk and a moving away from focusing on the individual offender towards actuarial considerations of aggregates (Feeley and Simon 1992). The recent incorporation of algorithms into this discourse is not merely a linear continuation of the new penology but an exponentiation which brings intrinsic novel challenges to criminal justice. As state actions against citizens are increasingly consolidated towards the preliminary stages of criminal investigations and preventive police-related settings, this provocation argues that legal protections must shift in the same direction to keep pace with technological developments.

Risk colonization: replacing ‘action’ with ‘behaviour’

Another reason to scrutinize and if necessary criticize the integration of automated risk prediction technologies into both law enforcement and pre-trial operations is that it may eventually impact criminal law beyond merely these two contexts. It may just be the beginning of data driven transformations affecting the whole criminal justice system, and it may lead to a shift in the focus of substantive criminal law away from concretely defined criminal offences to the more diffuse category of (algorithmically determined) criminally relevant behaviour and attitudes. The practice of measuring a person against minute data of her past behaviour is not novel in criminal law theory. It is reminiscent of criminal law theories of ‘Lifestyle-Guilt’ (Mezger 1938, 688) and ‘Life-Decision-Guilt’ (Bockelmann 1940, 145). These theories disassociated punishability and guilt from a single deliberate or negligent act and attached it to the inner nature of actors reflected in their past life choices. These approaches were popular in Germany during the Third Reich, and are particularly suited to autocratic rule due to their fluidity.

Already today predictive crime technologies and the data collected by such technologies do not remain solely within the realm of pre-trial decisions and policing, but have colonized court settings as well. In the U.S., even sentencing decisions are supported by data driven algorithmic analyses of an offender’s future behaviour. It is noteworthy that these analyses were initially developed only for the use in preventive law enforcement measures and only later migrated to a sentencing context (Angwin et al. 2016). A last step of this development may be the migration of algorithmic determinations to the establishment of criminal liability. It is worth noting that the possibility of supporting human rights judgements (Aletras et al. 2016), Supreme Court decisions (Islam et al. 2016), and civil litigation (Katz 2014) with data driven algorithms is already being sounded out. The potential future use of data driven predictions to determine criminal liability should thus not be discounted as a scenario too outlandish to prepare the legal system for.

Outlook: benign amusement and/or bitter reproach

It may well be that future generations, or at least their advantaged elites, will look back on our critical analyses of data driven government and smile benignly at us. Possibly with the same amusement we feel today, when we read about Plato’s warnings against the technology of writing (Plato, Phaedrus, 275A). A less-advantaged segment of future societies, however, comprised of individuals who were unable to profit from digital advancements, locked in negative algorithmic presumptions about themselves, may look back on us more reproachfully, asking why no safeguards for the rule of law and the PoI were pressed for and implemented.

This provocation does not aim to provide a definitive answer as to which of these scenarios is more likely to occur. Rather, it aims to prod the reader to think about whether and how traditional legal protections may be expanded to accommodate the rapid technological evolutions now changing the face of the criminal justice system. Particularly, it wants to prod the reader to consider if the PoI could be understood already today as a two-faced mechanism. A mechanism not just to ensure that a defendant in court will be presumed innocent until proven guilty. But a mechanism which also outside of a trial counters anticipatory data driven risk practices that preclude the possibility of a non-dangerous individual.

* Lucia Sommerer is a binational (German-American) legal scholar with focus on the intersection of criminal law and emerging technology. She is a Fellow at Yale Law School’s Information Society Project, a Master of Laws (LL.M.) candidate at Yale Law School, and a doctoral candidate at the chair for Criminal Law and Criminology at Göttingen University, Germany.

References

Aletras, Nikolaos, Dimitrios Tsarapatsanis, Daniel Preoţiuc-Pietro, and Vasileios Lampos. 2016. “Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective.” PeerJ Computer Science 2 (October): e93. doi: 10.7717/peerj-cs.93.

Angwin, Julia, Jeff Larson, Surya Mattu, and Lauren Kirchner. 2016. "Machine Bias: There’s Software Used Across the Country to Predict Future Criminals. And it’s Biased Against Blacks." ProPublica, May 23, 2016. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

Big Brother Watch. 2018. “Press Release: Police use Experian Marketing Data for AI Custody Decisions.” https://bigbrotherwatch.org.uk/all-media/police-use-experian-marketing-data-for-ai-custody-decisions/.

Bockelmann, Paul. 1940. Studien zum Täterstrafrech. Vol. 2. Berlin: de Gruyter.

Duff, Anthony. 2013. "Who Must Presume Whom to be Innocent of What?" The Netherlands Journal of Legal Philosophy 42:170-92.

Feeley, Malcom M., and Jonathan Simon. 1992. "The New Penology: Notes on the Emerging Strategy of Corrections and Its Implications" Criminology 30(4): 449-74.

Ferguson, Pamela R. 2016. "The Presumption of Innocence and its Role in the Criminal Process." Criminal Law Forum 27:131-58.

Islam, Mohammad Raihanul, K.S.M. Tozammel Hossain, Siddharth Krishnan, and Naren Ramakrishnan. 2016. "Inferring Multi-dimensional Ideal Points for US Supreme Court Justices." AAAI'16 Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, 4-12.

Katz, Pamela S. 2014. "Expert Robot: Using Artificial Intelligence to Assist Judges in Admitting Scientific Expert Testimony." Albany Law Journal of Science and Technology 24(1):1-47.

Mayson, Sandra G. 2018. "Dangerous Defendants." Yale Law Journal 127(3): 490-568.

Mezger, Edmund. 1938. "Die Straftat als Ganzes." Zeitschrift für die gesamte Strafrechtswissenschaft 57 (1): 675-701.

Oswald, Marion, Jamie Grace, Sheena Urwin, and Geoffrey Barnes. 2018. "Algorithmic risk assessment policing models: Lessons from the Durham HART model and 'Experimental' proportionality." Information & Communications Technology Law 27(2): 223-50. doi: 10.1080/13600834.2018.1458455.

Werth, Robert. 2018. "Theorizing the Performative Effects of Penal Risk Technologies: (Re)producing the Subject Who Must Be Dangerous." Social & Legal Studies Online First: 1-22. doi: 10.1177/0964663918773542.