Monique Mann

10 Automation

Abstract: This chapter explores the history and meaning of automation and its relevance in criminal justice. It is argued that there is a need to think critically about what 'automation' means and attend to the social context/s, socio-political systems, actors, and networks within them, as well as specific applications of 'automated' technologies to understand drivers and consequences at the intersection of science, technology, and society.

Keywords: automation, criminal justice, bias, discrimination

Introduction

The origins of automation stem from factory work and manufacturing with the objective of increasing efficiency, which over time has expanded to other areas such as health care, transportation, agriculture, construction, government administration, military, security, and criminal justice (Hitomi, 1994). There are various applications of automation in criminal justice. It is used, for example, to detect intrusions in computer systems, to predict recidivism, and to impose sentences or financial penalties. The use of automation in criminal justice is the focus of this contribution.

The term 'automaton' was coined by D. S. Harder while working for the General Motors Corporation in 1936, who later went on to establish the Automation Department at the Ford Motor Company (Hitomi, 1994). Therefore, automation has its origins in Fordist modes of production. At the time, Harder defined automation "as the transfer of work parts between the machines in a production process *without human operation*" (Hitomi, 1994: 122, emphasis added). However, this view of automation is deceptively simple and reductionistic as it does not consider wider socio-technical *assemblages* that shape 'automated' systems and processes (for an overview of the philosophical concept of 'assemblage' see Marcus and Saka, 2006).

'Automated' systems and processes may appear to operate independently of humans, but ultimately, they do not. 'Automation' is more complex than an absence of human influence, as it is constructed, enabled, and enacted by socio-political and human dimensions. These include decisions to collect information (or not), the analysis of information in certain ways for some purposes (and not others), the programming of algorithms to perform one function (over another), and they extend to how 'automated' outputs guide and influence human behavior (or not) (see Algorithm by Leese). Automation is always part of a wider *assemblage*, and as will be argued in this chapter, this concept is helpful to both understand and problematize 'automation' in criminal justice.

Open Access. © 2025 the author(s), published by De Gruyter. BY-NC-ND This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. https://doi.org/10.1515/9783111062037-011

Notwithstanding this critique, which is expanded on further below, historically, there are three main concepts of automation, as follows (developed by Diebold, Bright, and Druker, as cited in Hitomi, 1994: 122):

> 'Automation' may be considered the abbreviation of 'automatisation' or 'automatic operation'; alternatively automation is a combination of the Greek 'automotos' (meaning self-moving) and Latin –'ion' (meaning a state). 'Mechanization' is the replacement of human physical labour by machines, but the control of this machine operation is effected by human operators. However, 'automation' also replaces this control action by machines; that is 'automation' means the replacement of both human physical and mental activities by machines.

There is an important distinction, however, between automation and robotics (see Robots by Sandvik and Lintvedt). The IEEE Robotics and Automation Society Field of Interest Statement defines robotics as focusing on "systems incorporating sensors and actuators that operate autonomously or semi-autonomously in cooperation with humans" and defines automation (or research on it) with an emphasis on "*efficiency*, *productivity*, *quality*, and *reliability*, focusing on *systems that operate autonomously*, often in *structured environments* over *extended periods*, and on the explicit *structure of such environments.*" (Goldberg, 2012: 1, emphasis added). It is necessary, therefore, to consider the structure of environments which can be extended to include the social and political contexts in which 'automation' is both designed and deployed.

Relevance of automation in criminal justice

Automation in criminal justice has immense intellectual and criminological relevance as there has been an exponential increase in the collection, aggregation, and *automated* analysis of information and *automated* decision-making within the criminal justice system. This includes applications in digital forensics (Jarett and Choo, 2021), policing including predictive, pre-emptive, and intelligence-led approaches (Bennett Moses and Chan, 2014; 2016; Egbert and Mann, 2021), courts and sentencing, especially risk assessment (Završnik, 2019; 2020), and within prisons and carceral contexts (McKay, 2022). 'Big data' inform automation and automated decision-making, which has been termed "stochastic governance" (Sanders and Sheptycki, 2017). It has been argued that new forms of 'automatic justice' or 'automated justice' are challenging the traditional model of criminal justice and the discipline of criminology (Bowling et al., 2008; Chan and Bennett Moses, 2016; Marks et al., 2017; Brownsword and Harel, 2019). This is because divisions between surveillance, policing, adjudication, and punishment are being eroded with automation collapsing these processes and removing human actors, at least from obvious involvement (Marks et al., 2017; Završnik, 2019).

Automation may have the potential to improve the efficiency and accuracy of decision-making under certain circumstances or for some applications, but it can be used in ways that are harmful (Bennett Moses and Chan, 2014). This may include biases in datasets or the parameters of algorithms that underpin and facilitate automation targeting ‘risky’ individuals (Friedman and Nissenbaum, 1996), or automated technologies deployed against the most marginalized. These include populations that are more likely to be subject to disciplinary surveillance and control (Monahan, 2017) employed to ‘sort’ and subject them to differential (or discriminatory) treatment (Lyon, 2003). Automation in criminal justice is problematic as it involves the increased monitoring of people released from prison based on actuarial risk assessments (Harcourt, 2005). These processes are never neutral, and bias or errors may be difficult to contest with barriers to transparency and a false veil of objectivity provided by computerization (Friedman and Nissenbaum, 1996).

Conceptual complications of ‘automation’

There are different conceptual perspectives that can help understand, and complicate, ‘automation’ in criminal justice. Automated processes require input data and through analyzing information, automated decisions can be made, or automated actions completed. Therefore, the starting point that enables automation, is the collection of information via forms of surveillance (see Lyon: Surveillance), that may also be automated (Andrejevic, 2019). Significantly, surveillance practices and datasets converge into a “surveillant assemblage” (Haggerty and Ericson, 2000). Rather than involving independent datasets, discreet technologies or surveillance practices, Haggerty and Ericson (2000: 610) argue that “surveillance is driven by the desire to bring systems together, to combine practices and technologies and integrate them into a larger whole.” They also draw attention to the fact that surveillance is often directed on or towards the human body (for example, a face via facial recognition technology). According to Haggerty and Ericson (2000: 612–613) “the surveillant assemblage relies on machines to make and record discrete observations” by standardizing “the capture of flesh/information flows of the human body” and creating “data doubles.”

Automated decisions upon “data doubles” are occurring in many social institutions, not only criminal justice, but also, for example, welfare administration. Take, for instance, my own case study of the ‘RoboDebt’ automated debt raising scandal (Mann, 2020). This involved *automated* income matching of data about individuals held by and shared between various government agencies to *automatically* identify overpayment of benefits, and the *automated* issuance of debts *automatically* sent to welfare recipients. However, the program was driven by politics, rather than automated technologies (Mann, 2020).

Haggerty and Ericson’s (2000) “surveillant assemblage” and this example of automated debts connects to O’Malley’s (2010) concepts of “simulated justice” and “telemetric policing,” which take the “surveillant assemblage” further by theorizing shifts in the administration of ‘justice’ to the “monetisation of justice” (O’Malley, 2010: 795). O’Malley (2010) uses the example of *automatic* traffic fines, one that will be familiar to many readers: an offense is *automatically* detected by a camera, an individual (or its “data double”) is *automatically* identified by a number plate, connected to information stored in a roads traffic authority database (i. e., the "surveillant assemblage"), all of which leads to a fine being *automatically* posted. This occurs with little apparent involvement of criminal justice agents such as the police or judges. O'Malley (2010: 797) explains that the "nexus with money" is a central aspect of "simulated justice" as it enables financial penalties to be *automated.* Yet, humans are involved in conceptualizing, constructing, and authorizing the operation of such 'automated' systems, which are influenced by wider social and political dynamics such as the desire to improve road safety or raise revenue. So, are they truly 'automated'?

'Automated,' 'assisted,' or an 'assemblage'?

Facial recognition technology (FRT) is one technology that is often considered to be 'automated' (Mann and Smith, 2017). It is used here as an example to highlight the importance of the contexts and applications of technologies in criminal justice, and to question whether they are indeed 'automated.'

The main applications of FRT, at present, are one-to-one matching of biometric templates to automatically verify the identity of a person. This occurs at a border with access being granted if the biometric template of the individual physically standing at the smart gate matches the biometric template stored in their passport, and *provided that the individual, their passport and their 'data double' are deemed worthy of access.* This is a process of social sorting that reflects geo/political decisions about borders and citizenship (Lyon, 2003).

The second main application of FRT involves one-to-many searching. This could involve the CCTV footage, images from the internet, or those stored in other databases to identify someone unknown (Mann and Smith, 2017). The concepts of the 'surveillant assemblage,' 'simulated justice,' and 'telemetric policing' are relevant. This is because information drawn from the 'surveillant assemblage' is used to automatically identify an individual who may be deemed 'risky' (an assessment made from information held in databases and analyzed in automated ways; see Databases by Bellanova), or subject to an automated fine due to behavior remotely and automatically identified through surveillance.

It has been argued that facial recognition technology, when used for one-to-many searching, is not *automated per se,* as technologies and humans converge in hybrid ways and operate as part of a wider assemblage (Fussey et al., 2021). Fussey et al. (2021) argued that '*assisted*' FRT is a more appropriate term, given the ways in which FRT is used by police, for example by deciding about whether (or not) to stop someone that the FRT algorithm identifies. Police make discretionary decisions following identification by a FRT algorithm. Hence, one would rather speak of a hybrid interaction between police and the FRT algorithm (Fussey et al., 2021).

This is related to Egbert and Mann's (2021) critique of predictive policing technologies and their analysis of the ways in which police act upon the outputs of them (see Policing by Wilson; Prediction by Ķīlis, Gundhus, and Galis). For example, police may behave differently in areas that are predicted to be either high or low risk (Egbert and Mann, 2021). What is more, police may be more attentive in areas predicted to be 'high risk.' Furthermore, there are a range of factors that shape the inputs of predictive technologies, including the data that informs the algorithms, which are collected under certain social and political conditions, and are not neutral either. An example of such data are the types of crimes reported to or detected by police (Egbert and Mann, 2021).

The point of discussing the distinction between '*automated*' and '*assisted*' using the examples of FRT and predictive policing is to critique and problematize reductionistic conceptual binaries between 'automated' or otherwise. *All* technologies and their applications are mediated by a range of complex social, political, bureaucratic/institutional, and human factors within a wider assemblage. This has consequences for bias and discrimination (and not *only* in a technical sense), for example when 'automated' technologies in criminal justice tend to be deployed towards the most marginalized and those that already disproportionately have contact with the criminal justice system due to historical, social, and political reasons. Automated technologies do not exist in a social, political, or institutional vacuum.

Issues and concerns with automation in criminal justice

There are international attempts to regulate automation and automated decision-making through, for example, Article 22 of the European Union's General Data Protection Regulation (GDPR) that grants a right not to be subject to automated decisions and introduces a so-called 'right to explanation' (on limits to data protection and a 'right to explanation', see Wachter et al., 2017; Mann and Matzner, 2019). However, issues and concerns, especially given the potential human rights impacts, remain (Završnik, 2020). Examples of rights that may be impacted by automation include non-discrimination, especially when the logic of algorithms is to discriminate (Mann and Matzner, 2019).

Moreover, automated technologies operate at scale (one scale is, for example, population level in the case of traffic fines). There is a risk that due process safeguards may be undercut. An example includes the presumption of innocence and punishment for crimes committed predicted or pre-empted based on risk rather than committed with criminal conviction (see Garrett, 2013). A further example includes the right to a fair trial and being able to hear and understand the evidence, which may prove challenging with propriety algorithms. There are also a range of concerns such as biases that continue historical social trajectories and discriminate against women and People of Colour (Buolamwini and Gebru, 2018). It is often argued that it is necessary to have a 'human in the loop' to safeguard from the negative consequences of automation (see for example Lettieri et al., 2023). However, this may introduce other issues as 'assisted' or human decision-making are not free from bias either, as I have argued above and elsewhere (Egbert and Mann, 2021, see also Enarsson et al., 2021). This means that both human and assisted decision-making can be biased, with the solution to address bias. Therefore, this requires socio-political and structural solutions rather than only technological ones. Hence, the interaction/s between automation by computers (and datasets, analytic processes, and algorithms), socio-political systems, and humans as a complex assemblage demands greater study, theorizing, and empirical analysis.

Conclusion and future directions

Technology is developing rapidly, and the algorithms at the base of automation and automated decisions are becoming more complex (for example, 'Artificial Intelligence' or 'AI'). The number of sensors is increasing (e. g., at home, in the car, or on/in the body), a trend that will continue feeding the 'surveillant assemblage.' There will be more 'automated' 'telemetric policing' and 'simulated justice' opportunities. Automation will also operate in combination with and/or within other emerging policing technologies (such as forms of *robotic* policing). There are also highly controversial applications of automation that have been advocated for and may develop over time. For example, in the field of imprisonment this would involve automation for remote monitoring via surveillance systems and electronic monitoring devices such as ankle bracelets. These monitoring devices can include, for example, automated conducted energy devices and functionalities to 'automatically' immobilize (rather than fine) (Bagaric et al., 2018). Such projects demonstrate the importance of thinking critically about the applications for automation, and the impacts for individuals, their rights, and the principles of criminal justice.

Main takeaways

- The collection, assemblage, and analysis of information and associated 'automated' decisions or actions in criminal justice contexts is increasing. Concepts such as assemblage/s can help understand, critique, and problematize 'automation.'
- There is a need to think critically about what 'automation' means and attend to the social context/s, socio-political systems, actors, and networks within them, and specific applications of 'automated' technologies to understand drivers and consequences at the intersection of science, technology, and society.
- Automation presents risks of bias, discrimination, errors, and introduces challenges regarding transparency, fairness, justice, and recourse for those subject to it.
- There is an immediate need for strong and enforceable legal and regulatory frameworks to protect human rights and ensure due process. While some international models emerge (such as Article 22 of the GDPR or the more recent European Artificial Intelligence Act, but those have limits too), legal and regulatory issues still need discussion (see, for example, Wachter et al., 2017; Mann and Matzner, 2019).

Suggested reading

Andrejevic, M. (2019). Automating surveillance. *Surveillance and Society*, 17(1/2), 7 – 13.

Eubanks, V. (2018). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor.* New York City: St. Martin's Press.

Marks, A., Bowling, B., & Keenan, C. (2017). Automatic justice? Technology, crime and social control. In R. Brownsword, E. Scotford, & K. Yeung (eds.) *The Oxford Handbook of Law, Regulation and Technology* (pp. 705 – 730). Oxford: OUP.

References

Andrejevic, M. (2019). Automating surveillance. *Surveillance and Society*, 17(1/2), 7 – 13.

Bagaric, M., Hunter, D., & Wolf, G. (2018). Technological incarceration and the end of the prison crisis. *The Journal of Criminal Law and Criminology*, 108(1), 73 – 135.

Bennett Moses, L., & Chan, J. (2014). Using big data for legal and law enforcement decisions: Testing the new tools. *UNSW Law Journal*, 37(2), 643 – 678.

Bennett Moses, L., & Chan, J. (2016). Algorithmic prediction in policing: Assumptions, evaluation and accountability. *Policing and Society*, 8(7), 806 – 822.

Bowling, B., Marks, A., & Murphy, C. (2008). Crime control technologies: Towards an analytical framework and research agenda. In R. Brownsword & K. Yeung (eds.), *Regulating Technologies* (pp. 51 – 78). Oxford: Hart Publishing.

Brownsword, R., & Harel, A. (2019). Law, liberty and technology: Criminal justice in the context of smart machines. *International Journal of Law in Context*, 15(2), 107 – 125.

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, 81, 77 – 79.

Chan, J., & Bennett Moses, L. (2016). Is big data challenging criminology? *Theoretical Criminology*, 20(1), 21 – 39.

Egbert, S., & Mann, M. (2021). Discrimination in predictive policing: The dangerous myth of impartiality and the need for STS-Analysis. In A. Završnik & V. Badalic (eds.), *Automating Crime Prevention, Surveillance, and Military Operations* (pp. 25 – 46). Cham: Springer.

Enarsson, T., Enqvist, L., & Naarttijarvi, M (2021). Approaching the human in the loop – legal perspectives on hybrid human/algorithmic decision-making in three contexts. *Information & Communications Technology Law*, 1, 123 – 153.

Friedman, B., & Nissenbaum, H. (1996). Bias in computer systems. *ACM Transactions on Information Systems*, 14(3), 330 – 347.

Fussey, P., Davies, B., & Innes, M. (2021). 'Assisted' facial recognition and the reinvention of suspicion and discretion in digital policing. *British Journal of Criminology*, 61, 325 – 344.

Garrett, B. L. (2013). Big data and due process. *Cornell Law Review Online*, 99, 207 – 217.

Goldberg, K. (2012). What is automation? *IEEE Transactions on Automation Science and Engineering*, 9(1), 1 – 2.

Haggerty, K. D., & Ericson, R. V. (2000). The surveillant assemblage. *British Journal of Sociology*, 51(4), 605 – 622.

Harcourt, B. (2005). Against prediction: Sentencing, policing, and punishing in an actuarial age. *University of Chicago Public Law & Legal Theory Working Paper, No. 94.*

Hitomi, K. (1994). Automation – its concept and a short history. *Technovation*, 14(2), 121 – 128.

Jarett, A., & Choo, K. R. (2021). The impact of automation and artificial intelligence on digital forensics. *WIREs Forensic Science*, 3(6), 1 – 17.

Lettieri, N., Guarino, A., Zaccagnino, R., & Malandrino, D. (2023). Keeping judges in the loop: A human–machine collaboration strategy against the blind spots of AI in criminal justice. *Soft Computing*, 27, 11275 – 11293.

Lyon, D. (2003). Surveillance as social sorting: Computer codes and mobile bodies. In D. Lyon (ed.), *Surveillance as SOCIAL SORTING: Privacy, Risk and Digital Discrimination* (pp. 13 – 30). London and New York: Routledge.

Mann, M. (2020). Technological politics of automated welfare surveillance. *Global Perspectives*, 1(1), 1 – 12.

Mann, M., & Matzner, T. (2019). Challenging algorithmic profiling: The limits of data protection and anti-discrimination in responding to emergent discrimination. *Big Data and Society*, 6(2), 1 – 11.

Mann, M., & Smith, M. (2017). Automated facial recognition technology: Recent developments and strengthening oversight. *UNSW Law Journal*, 40(1), 121 – 145.

Marcus, G. E., & Saka, E. (2006). Assemblage. *Theory, Culture and Society*, 23(2 – 3), 101 – 109.

Marks, A., Bowling, B., & Keenan, C. (2017). Automatic justice? Technology, crime and social control. In R. Brownsword, E. Scotford & K. Yeung (eds.), *The Oxford Handbook of Law, Regulation and Technology* (pp. 705 – 730). Oxford: OUP.

McKay, C. (2022). The carceral automation: Digital prisons and technologies of detention. *International Journal of Crime Justice and Social Democracy*, 11(1), 100 – 119.

Monahan, T. (2017). Regulating belonging: Surveillance, inequality, and the cultural production of abjection. *Journal of Cultural Economy*, 10, 191 – 206.

O'Malley, P. (2010). Simulated justice: Risk, money and telemetric policing. *British Journal of Criminology*, 50, 795 – 807.

Sanders, C., & Sheptycki, J. (2017). Policing, crime and 'big data'; towards a critique of the moral economy of stochastic governance. *Crime, Law and Social Change*, 68, 1 – 15.

Wachter, S., Mittelstadt, B., and L. Floridi. (2017). Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation. *International Data Privacy Law*, 7(2), 76 – 99.

Završnik, A. (2019). Algorithmic justice: Algorithms and big data in criminal justice settings. *European Journal of Criminology*, 18(5), 623 – 773.

Završnik, A. (2020). Criminal justice, artificial intelligence systems, and human rights. *ERA Forum*, 20, 567 – 583.